Non-Oscillatory Third Order Fluctuation Splitting Schemes for Steady Scalar Conservation Laws

Matthew Hubbard

School of Computing, University of Leeds, Leeds, LS2 9JT, UK.

Abstract

This paper addresses the issue of constructing non-oscillatory, higher than second order, fluctuation splitting methods on unstructured triangular meshes. It highlights the reasons why existing approaches fail and proposes a procedure which can be applied to any high order fluctuation splitting scheme to impose positivity on it. Its success is demonstrated through application to a series of linear and nonlinear scalar problems, using a pseudo-time-stepping technique to reach steady state solutions on two-dimensional unstructured meshes.

Key words: Fluctuation Splitting, Hyperbolic Conservation Laws, Positivity, High Order Accuracy.
PACS: 65M99, 76M25

1   Introduction

The fluctuation splitting approach to approximating multidimensional systems of conservation laws has developed to a stage where it can be used reliably to produce accurate simulations of complex steady state fluid flow phenomena using unstructured meshes [12]. The most commonly used methods are second order accurate at the steady state, which is deemed accurate enough for simulating a wide range of flows and, in the presence of discontinuities, they are also required to avoid introducing unphysical oscillations in to the flow. Within the fluctuation splitting framework, the PSI scheme [13] (or one of its variants) has provided the basis for a range of methods which have successfully achieved these goals, particularly for the scalar case. It is a nonlinear upwind scheme

---

Email address: meh@comp.leeds.ac.uk (Matthew Hubbard).

which, when it is used to simulate scalar conservation laws, exhibits second order accuracy at the steady state for smooth solutions, guarantees positivity (even in the presence of discontinuities), and gives rapid convergence to the steady state, all without the necessity for additional artificial viscosity.

Generalisations of the PSI scheme have been developed for application to nonlinear systems of equations [19]. The most commonly used schemes for systems take the form of matrix distribution schemes. A nonlinear matrix distribution form of the PSI scheme exists, but it has so far proved more successful to reinterpret the PSI scheme as a nonlinear "blending" of two linear schemes, the non-oscillatory N scheme and the second order accurate LDA scheme, both of which have natural matrix forms which can be applied to nonlinear systems of equations [1]. Additional research, which led to the development of wave decomposition models (see [19] for a summary), has shown that by preconditioning the equations, an optimal decoupling can be achieved. This allows for scalar components, which take the form of advection equations, to be split off so that they can be approximated independently using the original PSI scheme [13], before recombining the components to give a conservative update to the solution. The work presented here demonstrates that it is possible to create a scheme which combines positivity with a higher order of accuracy (as well as higher accuracy) than the PSI scheme at the steady state. The discussion will be restricted to scalar equations, which can be applied directly to the simulation of passive transport phenomena or individual components of a more complex system, though their use in the latter is left to future research, since it is not yet clear what their ideal extension to nonlinear systems should be.

More recently, research has focused on the development of more accurate fluctuation splitting methods, for both steady state and time-dependent problems. This has led to the creation of a number of third and higher order accurate schemes, each one based on a standard approach to constructing a higher order representation of the dependent variables.

The first third order fluctuation splitting scheme to be developed was that of Caraeni *et al.* [10,11], who created a quadratic representation within each mesh cell via the reconstruction of local gradients of the dependent variables at the mesh nodes, obtained using the surrounding data. The resulting fluctuation was distributed using the non-positive LDA scheme [13] so, although the results shown for smooth viscous fluid flow and turbulence transport were excellent, unphysical oscillations still occur in less smooth situations.

Abgrall, along with Roe [8] and Mezine and Andrianov [4,7], proposed a similar idea, but constructed the higher order fluctuation using extra information about the dependent variable stored at the additional nodes created by a uniform global subdivision of the mesh. The solution is stored and updated

at all of the submesh nodes and the distribution of the fluctuations is carried out on the subtriangles. The proposed third (and higher) order schemes are *almost* non-oscillatory.

A third, less successful, approach was proposed by Hubbard and Laird [17], who obtained higher order by extending the stencil of the distribution of the second order fluctuations. It was highly sensitive to the mesh structure, and lacked robustness, but the underlying concept (of extending the stencil) remains valid as an approach to achieving higher order accuracy if it is used instead to construct a higher order representation of the dependent variable.

This work will consider the first two of these techniques for constructing a continuous piecewise quadratic interpolant, which leads to third order methods. This higher order interpolant is used to construct a more accurate fluctuation which, if distributed completely using bounded distribution coefficients, gives a high order method [1,6]. It is straightforward to generalise the concepts to higher than third order, though the implementation will be more complicated.

Each approach has achieved higher than second order accuracy for the scalar advection equation, and that of Caraeni *et al.* has already shown a great deal of promise in more practical situations. However, none of them has yet combined this with positivity, except with the aid of the Flux-Corrected Transport (FCT) algorithm [18,24] used with second order schemes in [14,17]. This paper will propose a method for imposing positivity on these high order schemes which avoids the use of any post-processing techniques such as FCT and has the additional benefit (not discussed here in detail) of providing a simple framework for creating a positive p-refinement algorithm for fluctuation splitting.

The research presented considers the scalar advection equation and a form of the inviscid Burgers' equation, approximated on two-dimensional, unstructured, triangular meshes. The fundamentals of first and second order fluctuation splitting schemes will be summarised in Section 2, after which the existing high order methods, of Caraeni *et al.* [10] and Abgrall and Roe [8], will be outlined in Section 3, along with the reasons underlying their lack of positivity. The discussion will suggest a modification which can be applied to any of these approaches in the steady state case to impose positivity on the scheme, and this is presented in detail in Section 4. A series of results for standard linear and nonlinear test cases will be given in Sections 5 and 6 to illustrate the effectiveness of the new approach at removing unwanted oscillations without unduly affecting the underlying scheme's accuracy. Finally, some conclusions are drawn in Section 7, alongside a brief description of possible directions for future work.

2 Fluctuation Splitting

Consider the two-dimensional scalar conservation law given by

$$u_t + f_x + g_y = 0 \qquad \text{or} \qquad u_t + \vec{\lambda} \cdot \vec{\nabla} u = 0 \qquad (1)$$

on a domain Ω, with $u(x, y, t) = h(x, y, t)$ imposed on the inflow part of the boundary $\partial\Omega$. $\vec{\lambda} = \left(\frac{\partial f}{\partial u}, \frac{\partial g}{\partial u}\right)^T$ defines the advection velocity associated with the conservation law. This equation has an associated fluctuation, assumed here to be calculated over a triangular mesh cell $\triangle$ and given by

$$\phi = -\int_{\triangle} \vec{\lambda} \cdot \vec{\nabla} u \, d\Omega = \oint_{\partial\triangle} u \, \vec{\lambda} \cdot \vec{n} \, d\Gamma \,, \qquad (2)$$

in which $\vec{n}$ represents the inward pointing unit normal to the cell boundary. When u is assumed to have a piecewise linear continuous representation with values stored at the mesh nodes, the discrete counterpart of ϕ is evaluated using an appropriate (conservative) linearisation [13]. Ideally, this allows the integration in Equation (2) to be carried out exactly, giving

$$\phi = -S_\triangle \, \tilde{\vec{\lambda}} \cdot \vec{\nabla} u = -\frac{1}{2} \sum_{i \in \triangle} u_i \, \tilde{\vec{\lambda}} \cdot \vec{n}_i \,, \qquad (3)$$

where $S_\triangle$ is the cell area and the symbol ˜ indicates an appropriately linearised quantity, in this case the local advection velocity. The index i loops over the vertices of $\triangle$, and $\vec{n}_i$ is the inward unit normal to the ith edge (opposite the ith vertex) multiplied by the length of that edge. This linearisation is straightforward in the special cases of divergence-free linear advection and Burgers' equation, the examples which will be used later in this paper. In both cases it equates to taking the arithmetic mean of the cell vertex values.

A simple forward Euler discretisation of the time derivative leads to an iterative update of the nodal solution values which is generally written as

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{S_i} \sum_{j \in \cup\triangle_i} \alpha_i^j \, \phi_j \,, \qquad (4)$$

where Δt is the time-step, S_i is the area of the median dual cell corresponding to node i (one third of the total area of the triangles with a vertex at i), α_i^j is the distribution coefficient which indicates the appropriate proportion of the fluctuation ϕ_j to be sent from cell j to node i, and $\cup\triangle_i$ represents the set of cells with vertices at node i. Conservation is assured as long as

$$\sum_{i \in \triangle_j} \alpha_i^j = 1 \qquad \forall j \,, \qquad (5)$$

4

where $\triangle_j$ represents the set of nodes at the vertices of cell j, *i.e.* the whole of each fluctuation is sent to the nodes. Note that the distribution has been restricted here so that a cell can only make contributions to nodes at its own vertices, allowing the scheme to be implemented very efficiently. The time derivative term in this construction is included here purely as a device for iterating to the steady state, but its presence would be necessary for time-dependent problems, in which case it must be integrated in a manner consistent with the underlying representation of u if the order of accuracy of the steady state approach is to be maintained [4,10].

2.1 The N Scheme

The most successful attempts to impose positivity on higher than first order fluctuation splitting schemes rely heavily on the linear first order, positive fluctuation distribution scheme known as the N scheme [13], which is defined as follows:

1 For each triangle, locate the downstream vertices, *i.e.* those for which $\tilde{\vec{\lambda}} \cdot \vec{n}_i > 0$, where $\vec{n}_i$ is the inward pointing normal to the edge opposite vertex i.

2a If a triangle has a single downstream vertex, node i_1 say (*cf.* Figure 1), then that node receives the whole fluctuation ϕ, so

$$ u_{i_1} \; \rightarrow \; u_{i_1} + \frac{\Delta t}{S_{i_1}} \, \phi \,, \tag{6} $$

while the values of u at the other two vertices remain unchanged.

2b Otherwise, the triangle has two downstream vertices, i_1 and i_2 say (*cf.* Figure 1), and the fluctuation ϕ is divided between these two nodes so that

$$ u_{i_1} \; \rightarrow \; u_{i_1} + \frac{\Delta t}{S_{i_1}} \, \phi_{i_1} \qquad u_{i_2} \; \rightarrow \; u_{i_2} + \frac{\Delta t}{S_{i_2}} \, \phi_{i_2} \,, \tag{7} $$

where

$$ \phi_{i_1} \; = \; -\frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_{i_1} (u_{i_1} - u_{i_3}) \qquad \phi_{i_2} \; = \; -\frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_{i_2} (u_{i_2} - u_{i_3}) \,, \tag{8} $$

in which i_3 denotes the remaining (upstream) vertex of the triangle. It is easily shown that $\phi_{i_1} + \phi_{i_2} = \phi$ (for conservation).

The distribution coefficients, α_i^j in Equation (4), can be derived easily from Equations (6)–(8). The resulting scheme is clearly locally, and hence globally, positive so the iteration given by (4) is conditionally stable. The appropriate

restriction on the time-step at node i is given by

$$\Delta t \;\leq\; \frac{S_i}{\sum_{j \in \cup \triangle_i} \max\left(0, \frac{1}{2}\tilde{\vec{\lambda}} \cdot \vec{n}_i^j\right)} \,.$$
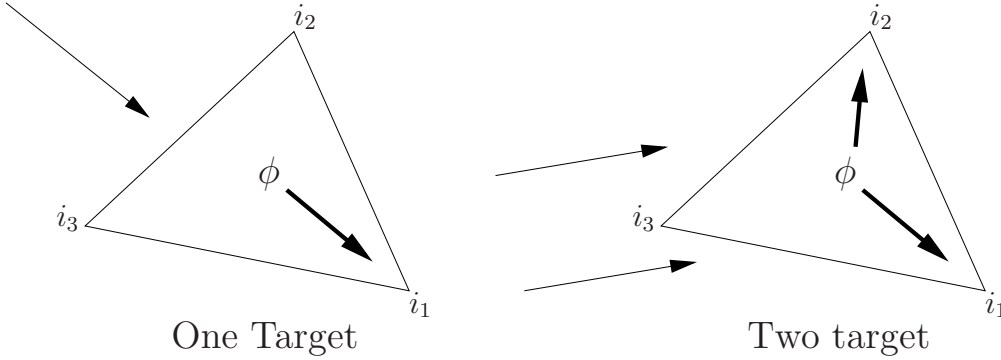(9)

One Target   Two target

Fig. 1. The splitting of the fluctuation distribution in to one and two target cases.

Results: The test case used here to illustrate the properties of each scheme consists of advection in a circle, with velocity $\vec{\lambda} = (y, -x)^T$ and over the domain $[-1,1] \times [0,1]$, of the initial profile given by

$$u(x,y,0) \;=\; \begin{cases} G(x) & \text{for } -0.65 \leq x \leq -0.35\,, \ y = 0, \\ 0 & \text{otherwise.} \end{cases}$$
(10)

This is also imposed as the boundary function $h(x,y,t)$ on the inflow boundaries of the domain while the experiment is run to steady state. The exact solution is $u(x,y) = G(r)$ for $0.35 \leq r = \sqrt{x^2 + y^2} \leq 0.65$, with $u(x,y) = 0$ elsewhere. In this case $\vec{\nabla} \cdot \vec{\lambda} = 0$, so the advection equation and the conservation law are equivalent and the conservative linearisation is simple [13].

Results are shown in Figure 3 for a uniform but genuinely unstructured triangular mesh consisting of 3806 nodes and 7370 cells ($dx \approx 0.025$), shown at the top of Figure 2. The two cases used are defined by

Case A: $G(r) = 1$, which illustrates the positivity of the scheme;
Case B: $G(r) = \cos^2(10\pi(r + 0.5)/3)$, which is more appropriate for determining the scheme's ability to maintain a smooth peak without artificially steepening the profile.

The solution profile along the outflow boundary at $y = 0$ would ideally reflect the profile at inflow exactly, but both here show the significant level of numerical diffusion incurred. This is unsurprising as the method is easily shown to be at best first order accurate [13].
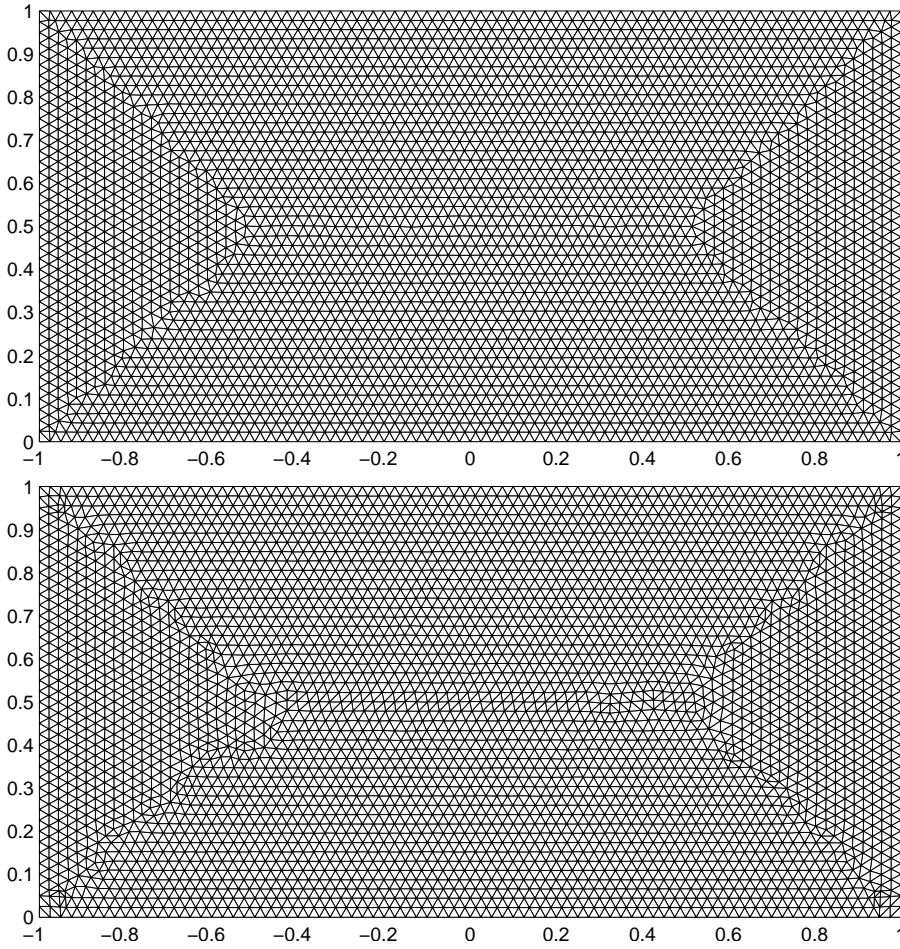
Fig. 2. The meshes used for the advection equation results obtained with the N, PSI and gradient recovery schemes (top) and the submesh reconstruction schemes (bottom).

2.2 The PSI Scheme

The nonlinear PSI scheme, devised by Struijs [23] and formulated algebraically by Sidilkover and Roe [22], is the most commonly used of the second order positive fluctuation splitting schemes, and is easily defined once the N scheme has been described.

Given that the contribution made by cell j to node i by the N scheme can be written as $(\phi_i^j)^N = (\alpha_i^j)^N \phi_j$, where ϕ_j is the fluctuation in cell j (see (3)), the contributions due to the PSI scheme can be defined as follows:

$$(\phi_i^j)^{PSI} \;=\; \frac{[(\alpha_i^j)^N]^+}{\sum_{k \in \triangle_j} [(\alpha_k^j)^N]^+} \phi_j \;=\; (\alpha_i^j)^{PSI} \phi_j \,, \tag{11}$$

in which $[\;\;]^+$ denotes the positive part of the quantity within the square brackets. This scheme has a number of notable properties (for all nodes i and
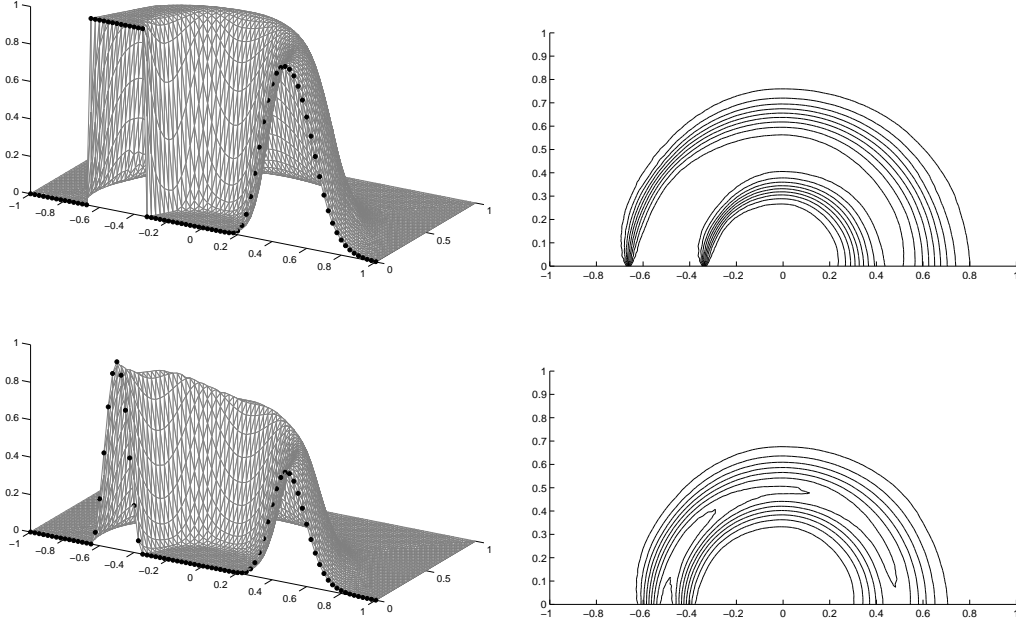
7

Fig. 3. The N scheme applied to Test Case A (top) and Test Case B (bottom).

cells j):

- The scheme is conservative because

$$\sum_{k\in\triangle_j} (\alpha_k^j)^{PSI} \; = \; \sum_{k\in\triangle_j} (\alpha_k^j)^N \; = \; 1 \, . \tag{12}$$

- $(\alpha_i^j)^{PSI} (\alpha_i^j)^N \geq 0$, so the scheme is positive for an appropriate range of time-steps.
- $|(\alpha_i^j)^{PSI}| \leq |(\alpha_i^j)^N|$, so the limit on the time-step given by (9) is sufficient to maintain this positivity.
- $(\alpha_i^j)^{PSI} \in [0,1]$ is bounded, so the order of accuracy of the steady state scheme is equivalent to the order of the error in the representation of ϕ (in this case second order) [1,6]. This property is often referred to as linearity preservation.

Results: The results for the PSI scheme are shown in Figure 4 using the same mesh as before (shown at the top of Figure 2). The improvement in accuracy over the N scheme is clear, as is the lack of oscillations. This is confirmed by the more comprehensive comparison of results supplied in Section 5. However, the outflow profile still does not reflect the inflow profile exactly (the square wave has clearly had its "corners" smoothed while the smooth profile has lost height on its peak value) and could be improved by incorporating a more accurate representation of the dependent variable in to the method.
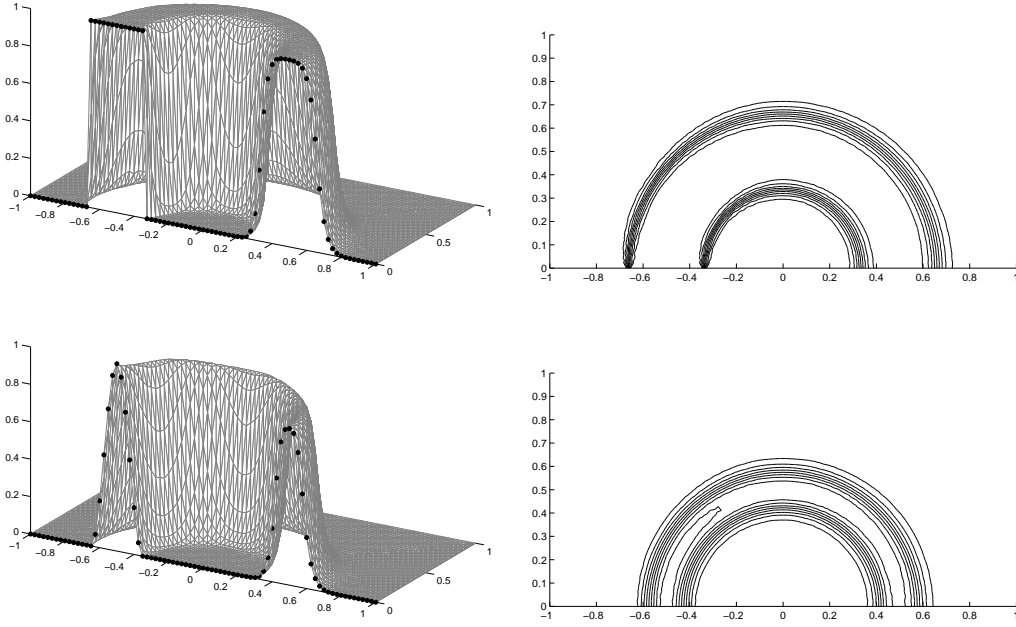
Fig. 4. The PSI scheme applied to Test Case A (top) and Test Case B (bottom).

3 Higher Order Methods

A range of techniques exist for extending numerical algorithms to higher order accuracy. Two which have been successfully applied to fluctuation splitting will be described here as a precursor to the introduction of the proposed approach for imposing positivity.

3.1 Submesh Reconstruction

The scheme of Abgrall and Roe [8] is based on a simple generalisation which differs mainly from the schemes described in Section 2 in that the fluctuation ϕ is approximated using a higher order polynomial representation of the dependent variable. The procedure can be extended to arbitrary orders of accuracy and to three space dimensions, but only the two-dimensional, third order case will be considered here.

In this situation the dependent variable u is taken to be a continuous piece-wise quadratic function with the unknowns stored at the nodes of a globally refined mesh, created by subdividing each triangular cell in to four congruent subcells (see the left hand diagram in Figure 5). This makes it possible to construct a unique local quadratic interpolating polynomial on each cell of the original mesh. The resulting fluctuation (2) can be evaluated on any of the subcells, using an appropriate quadrature rule to ensure that the integral

9

is evaluated to a high enough degree of accuracy. Exact integration will be assumed throughout this work, as it is required in the following analysis of each scheme's positivity, but this is not necessary to achieve the desired order of accuracy for the overall method. In fact, for a kth order method it suffices to evaluate the fluctuation exactly with respect to a $(k-1)^{\text{th}}$ degree polynomial representation of the flux and distribute this in a linearity preserving manner [8]. The resulting fluctuations, now calculated within each subcell, are then distributed to the nodes of the *refined* mesh.
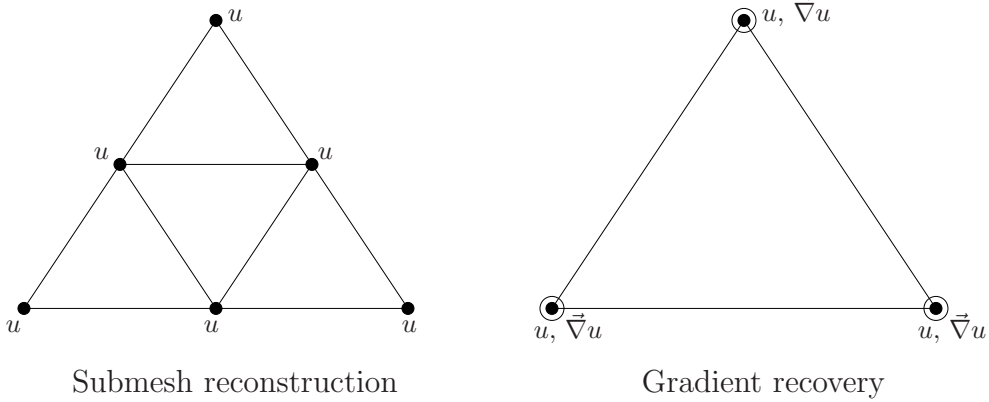

Submesh reconstruction   Gradient recovery

Fig. 5. The storage of solution information for the quadratic reconstructions of Abgrall and Roe (left) and Caraeni *et al.* (right). A filled circle indicates that a value of u is stored there and an unfilled circle indicates that a value of $\vec{\nabla}u$ is calculated there.

The method proposed in [8] is based on the contributions given by

$$(\phi_i^j)^{HO} = \frac{[(\alpha_i^j)^{N'}]^+}{\sum_{k\in\triangle_j}[(\alpha_k^j)^{N'}]^+} \, \phi_j^{HO} = (\alpha_i^j)^{HO} \, \phi_j^{HO} \,, \qquad (13)$$

in which ϕ_j^{HO} is calculated by evaluating (2) exactly according to the high order (HO) representation of u. The indices in (13), and for all schemes based on this submesh reconstruction approach, now represent vertex i (or k in the summation) of *subcell j*. As a consequence, the distribution of the fluctuation within each subcell remains local to that subcell, although the scheme is not completely local because the reconstruction of the polynomial within the subcell takes information from the rest of the full mesh cell.

The fluctuations defined by (13) are closely related to those of the PSI scheme (11) in form, but it is important to note that the N scheme-like distribution coefficients, denoted by the superscript N', have a slightly different definition, *i.e.*

$$(\alpha_i^j)^{N'} = (\phi_i^j)^N/\phi_j^{HO} \,, \qquad (14)$$

in which $(\phi_i^j)^N$ is the contribution made to node i when the N scheme is applied to the fluctuation created by integrating a linear representation of the dependent variable on subcell j. Furthermore, for reasons which will be

10

discussed in more detail in Section 3.3, the distribution coefficients which are actually used are

$$(\alpha_i^j)^{HO} \;=\; \frac{[(\alpha_i^j)^{N'}]^+ + \epsilon}{\sum_{k \in \triangle_j}([(\alpha_k^j)^{N'}]^+ + \epsilon)}\,, \tag{15}$$

where ϵ is some small free parameter.

Results: The results for the Abgrall-Roe scheme are presented in Figure 6, obtained using the mesh shown at the bottom of Figure 2, which has been constructed by regular subdivision of a 984 node, 1846 cell uniform but unstructured mesh (giving 3813 nodes and 7384 subcells with $dx \approx 0.025$, a similar number of unknowns as used in the other schemes) so that the subcells are similar in size to the cells of the mesh used for the other methods. An improvement in accuracy is apparent in the profile at outflow of both test cases (and is illustrated more clearly by the error measures given in Tables 1 and 2 in Section 5). However, this is at the expense of small oscillations just visible at the discontinuities close to the inflow boundary in Test Case A (again, see Tables 1 and 2).
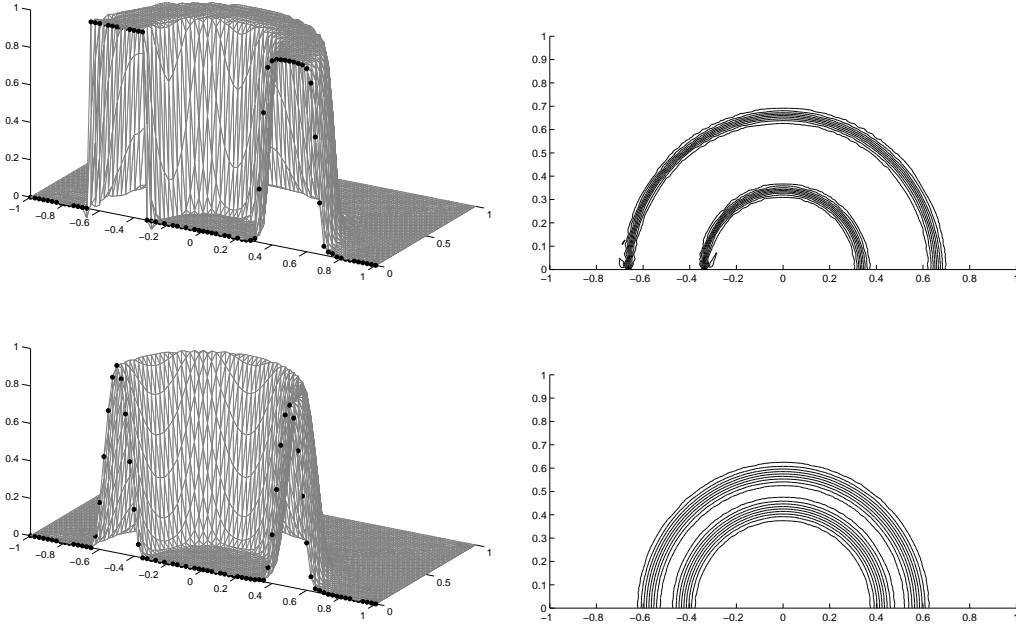

Fig. 6. The Abgrall-Roe scheme applied to Test Case A (top) and Test Case B (bottom).

3.2 Gradient Reconstruction

One alternative to subdividing the mesh to provide the additional degrees of freedom necessary to reconstruct quadratic polynomials is to recover solution

gradients at the mesh nodes and use these to obtain a quadratic interpolant. This approach was taken by Caraeni and his collaborators [10,11]. The first stage of the procedure used approximates the cell gradients using the Green-Gauss formulation, given by

$$\vec{\nabla} u_j = \frac{1}{S_{\triangle_j}} \oint_{\partial \triangle_j} u \, \vec{n} \, d\Gamma \,, \tag{16}$$

approximating the integral along each cell edge using the trapezoidal rule. These are then used to obtain an approximation to the nodal solution gradients from

$$\vec{\nabla} u_i = \frac{1}{\sum_{j \in \cup \triangle_i} |S_{\triangle_j}|^{-1}} \sum_{j \in \cup \triangle_i} |S_{\triangle_j}|^{-1} \vec{\nabla} u_j \,. \tag{17}$$

This is second order accurate on uniform meshes but loses this property on the unstructured meshes shown in Figure 2, most seriously along the "diagonals" which clearly appear in the mesh structure, emanating from each corner of the domain.

When quadrature is used to evaluate the fluctuations (2), the midpoint values of u are required (in the linear, third order case), and these are defined here by

$$u_{mid} = \frac{u_{i_1} + u_{i_2}}{2} + \frac{\vec{\nabla} u_{i_1} - \vec{\nabla} u_{i_2}}{8} \cdot (\vec{x}_{i_2} - \vec{x}_{i_1}) \,, \tag{18}$$

where i_1 and i_2 denote the vertices at each end of the mesh edge. This is used to calculate a high order fluctuation within the cell, which is then distributed using the LDA scheme. As with the other schemes, the distribution of the fluctuation within each cell remains local to that cell. However, the reconstruction of the gradient used in (17) takes information from the surrounding cells so the stencil of the scheme is not local, and for even higher order accuracy the stencil would need to be extended further in order to provide approximations to higher derivatives. The submesh reconstruction approach encounters a similar issue, in the sense that the stencil of the scheme used to update a given node extends beyond the adjacent subcells, but the structure inherent in the subdivision ensures that the distribution of the fluctuation within a full mesh cell only depends on information within that cell, whatever the order of accuracy required.

The gradient reconstruction scheme is not positive (and was never intended to be) but does achieve higher order accuracy, and gives excellent approximations to a range of smooth solutions [9–11].

Results: The results obtained using Caraeni's scheme are shown in Figure 7 using the same mesh as was used for the N and PSI schemes (the top mesh of Figure 2). The improvement in accuracy is clear in the modelling of the smooth profile but the effects of its lack of positivity are equally apparent

12

in the test case involving the discontinuous profile. This is confirmed by the evidence presented in Tables 1 and 2 of Section 5.
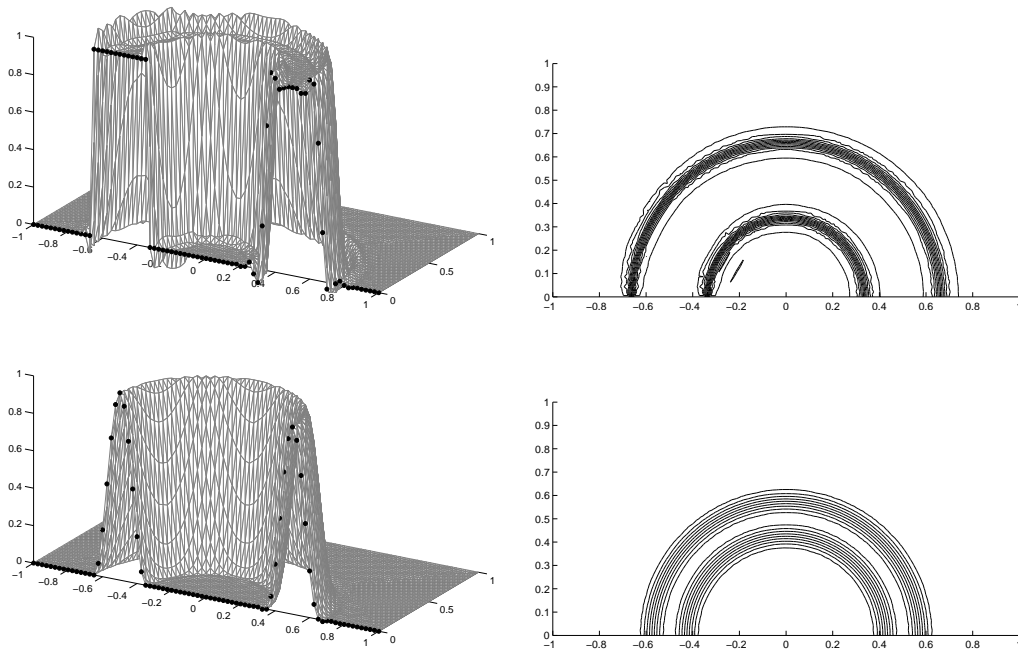

Fig. 7. The Caraeni scheme applied to Test Case A (top) and Test Case B (bottom).

3.3   The Problem

In the following discussion ϕ_j^{HO} will be used to denote the fluctuation within a mesh cell or, for any scheme based on uniform subdivision of the mesh, subcell indexed by j and evaluated from the high order (quadratic) interpolant. ϕ_j^{LO} will denote the fluctuation within the same (sub)cell but evaluated from the low order (linear) interpolant. The analysis follows that given in [15] and closely resembles that presented in independent work by Ricchiuto *et al.* [21] and Abgrall [2].

Three situations are worthy of discussion, none of which arises in the PSI scheme because, in that case, $\phi_j^{HO} \equiv \phi_j^{LO}$.

1 $\phi_j^{HO} \phi_j^{LO} < 0$ for some (sub)cells j, *i.e.* the high and low order fluctuations can have different signs. This is most damaging when the N scheme gives only non-negative distribution coefficients $(\alpha_i^j)^N$. In such cases the procedure used by the Abgrall-Roe scheme (13) which, in the second order case, imposes linearity preservation on the N scheme and leads to the PSI scheme, now gives zero distribution coefficients for every vertex of the (sub)cell, clearly violating conservation. The modified form given in (15) avoids this loss of conservation, but only by reverting to a central discreti-

13

sation, $\alpha_i^j = \frac{1}{3}$, $\forall i \in \triangle_j$, in such situations, which is not positive. The resulting scheme is also not continuous, which can present problems when attempting to converge to steady state solutions.

2 $|\phi_j^{HO}| \gg |\phi_j^{LO}|$ for some (sub)cells j, even when $\phi_j^{HO} \phi_j^{LO} \geq 0$, *i.e.* the magnitude of the high order fluctuation can be much higher than that of the low order fluctuation, even when they have the same sign. Since the ratio of the two fluctuations is unbounded this can seriously restrict the condition (9) required for the time-stepping procedure to remain positive, and hence the speed with which a positive steady state approximation can be found (if one even exists).

3 The most troublesome case of all, which occurs in many (sub)cells in Test Case A, is when ϕ_j^{HO} is nonzero in a (sub)cell for which $u_{i_1} = u_{i_2} = u_{i_3}$. In such situations it is impossible to distribute ϕ_j^{HO} to the vertices of (sub)cell j in a conservative manner while maintaining local positivity. Global positivity cannot therefore be achieved without prior knowledge of the contributions from elsewhere. The consequence of this is that it is not possible to construct a positive fluctuation distribution scheme which is conservative and higher than second order accurate if the distribution of the fluctuation in a (sub)cell is restricted to be only to the vertices of that (sub)cell. Thus one of the main design criteria used in the development of fluctuation splitting methods cannot be enforced.

Remark: At this stage of the development of a high order non-oscillatory *finite volume* scheme a limiter would be introduced to combine the low order positive scheme with the high order non-positive scheme. Applying such an approach here (at least in the manner traditionally employed in finite volume schemes) would lead to an inconsistency in the calculations of fluctuations in neighbouring (sub)cells, in that the edge contributions in the boundary integrals in (2) would no longer necessarily cancel at internal edges, leading to a non-conservative scheme. This is an issue which this work seeks to address, though it is worth noting that Abgrall and Barth [5] have demonstrated that this loss of conservation may not always be a problem in practice. In particular, if discontinuities in the representation could be confined to smooth regions of the flow then the lack of conservation should not have a detrimental effect on the quality of the solution. This possibility will not be explored here, but the results produced suggest that a scheme of this form might be constructed in the future (see Figure 16 and the accompanying discussion in Section 5).

4 Suppressing the Oscillations

A number of techniques exist which can be used to remove unphysical oscillations from computational simulations of hyperbolic conservation laws. For

example, the Flux-Corrected Transport (FCT) approach [18,24] is widely used and has been applied within the fluctuation distribution framework [14,16,17]. In this work though, the intention is to incorporate local positivity within the method itself, rather than applying a post-processing step once low and high order updates have already been calculated.

It has been observed that, in the submesh distribution framework of Abgrall and Roe [8], it is always possible to distribute the high order fluctuation of the full mesh cell J, $\phi_J^{HO} = \sum_{\triangle_k \in \triangle_J} \phi_k^{HO}$, to the six subvertices of that cell in a manner which is both conservative and locally positive [2,15,20]. However, this observation has yet to yield a scheme which combines positivity with high order accuracy. The schemes presented in [15] strove for higher order accuracy, and although they improved on the PSI scheme in some cases and were free of oscillations, they couldn't be proved to be positive. That technique is also essentially still a post-processing step, can only be applied if the submesh reconstruction approach is used to obtain higher order accuracy, and doesn't offer a straightforward generalisation to higher orders of accuracy. A second, considerably more successful and flexible approach, will be presented here.

4.1   Limiting the Interpolant

The modification presented in [15] addresses the problem of distributing the high order fluctuation in a positivity-preserving manner by, essentially, extending the stencil. An alternative approach is to modify the interpolant (and hence the fluctuation) in a manner which allows a positive distribution scheme within the existing framework, *i.e.* one where the fluctuation in a (sub)cell is only distributed to that (sub)cell's vertices. The linear interpolant used in the second order scheme has already led to positive schemes of this type, *e.g.* N, PSI, so the challenge is to find a procedure (similar in aim to flux/slope limiting in finite volume schemes) which can be used to add a "limited" amount of a high order correction to a low order scheme to improve its accuracy while maintaining positivity.

The approach to approximating Equation (1) taken here consists of two stages. First it is necessary to apply a simple adjustment to the interpolant, so that it is possible to distribute the resulting fluctuation within each triangular (sub)cell to its vertices in a locally positive manner. Once the fluctuation due to this adjusted interpolant has been calculated, a second modification needs to be made, this time to the distribution scheme.

In order to determine the adjustment which must be made to the interpolant, consider $\bar{u}(\vec{x})$, the linear interpolant of the values of the dependent variable u at the vertices of a given triangular (sub)cell, and $\delta u(\vec{x})$ the high order

15

correction to the interpolant over that triangle, for which $\bar{u}(\vec{x}) + \delta u(\vec{x})$ is a high order representation of the data within that triangle (independent of the method used to obtain it).

The fluctuation due to the high order interpolant is given by

$$\oint_{\partial\triangle} u\,\vec{\lambda}\cdot\vec{n}\;d\Gamma = \sum_{edges} \int_{i_1}^{i_2} u\,\vec{\lambda}\cdot\vec{n}\;d\Gamma$$
$$= \sum_{edges} \left[\int_{i_1}^{i_2} \bar{u}\,\vec{\lambda}\cdot\vec{n}\;d\Gamma + \int_{i_1}^{i_2} \delta u\,\vec{\lambda}\cdot\vec{n}\;d\Gamma \right] \qquad (19)$$

in which i_1 and i_2 represent successive vertices around the triangular mesh (sub)cell (in an anticlockwise sense). The first term on the right hand side can easily be distributed in a positive manner using, for example, the N or PSI scheme. The second is more problematic so here it is replaced by a "limited" high order correction term which should return the high order interpolant whenever possible, but restrict the correction where this is necessary to achieve positivity.

It is simple to show that if the high order correction along each edge $i_1 i_2$ of the triangle is limited to give a modified high order correction $\delta u'(\vec{x})$ along that edge which satisfies

$$|\delta u'_{i_1 i_2}(\vec{x})| \;\leq\; C|u_{i_1} - u_{i_2}| \qquad \forall\,\vec{x} = \mu\,\vec{x}_{i_1} + (1-\mu)\,\vec{x}_{i_2}\,, \;\; 0 \leq \mu \leq 1\,, \quad (20)$$

for some finite constant $C \geq 0$, then, subject to an appropriate restriction on the time-step in (4), it is possible to distribute the fluctuation (2) due to the modified interpolant $u'(\vec{x}) = \bar{u}(\vec{x}) + \delta u'(\vec{x})$ to the vertices of that mesh (sub)cell in a locally positive manner.

The proof is straightforward once the fluctuation due to the "limited" interpolant $u'(\vec{x})$ is written in the form

$$\oint_{\partial\triangle} u'\,\vec{\lambda}\cdot\vec{n}\;d\Gamma = \sum_{edges} \int_{i_1}^{i_2} u'\,\vec{\lambda}\cdot\vec{n}\;d\Gamma$$
$$= \sum_{edges} \left[\int_{i_1}^{i_2} \bar{u}\,\vec{\lambda}\cdot\vec{n}\;d\Gamma + \int_{i_1}^{i_2} \delta u'\,\vec{\lambda}\cdot\vec{n}\;d\Gamma \right]. \qquad (21)$$

By defining a function $\alpha_{i_1 i_2}(\vec{x})$ for each edge, which is given by

$$\delta u'_{i_1 i_2}(\vec{x}) \;=\; \alpha_{i_1 i_2}(\vec{x})(u_{i_1} - u_{i_2})\,, \qquad (22)$$

it follows immediately that imposing (20) on this edge leads to $|\alpha_{i_1 i_2}(\vec{x})| \leq C$

for any $\vec{x}$ on the edge, and then from (21) that

$$\oint_{\partial\triangle} u' \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \;\; = \;\; \sum_{edges} \left[\int_{i_1}^{i_2} \alpha_{i_1 i_2}(u_{i_1} - u_{i_2}) \, \vec{\lambda} \cdot \vec{n} \; d\Gamma + \int_{i_1}^{i_2} \bar{u} \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \right] . \quad (23)$$

Clearly, given that α is bounded, this additional component can be distributed in a positive manner (for an appropriate limit on the time-step) by sending the component corresponding to the integral along edge $i_1 i_2$ to vertex i_1 or i_2 depending on the sign of the integral.

The following discussion suggests a particular form for the distribution which makes use of the notion of upwinding, as introduced in the description of the N scheme, and leads to the modified schemes proposed here. Noting that

$$(u_j - u_k) \;\; \equiv \;\; (u_i - u_k) - (u_i - u_j) \quad (24)$$

and that, from (3), the low order fluctuation on the mesh (sub)cell is

$$\phi^{LO} \;\; = \;\; -\frac{1}{2} \sum_{i \in \triangle} u_i \, \tilde{\vec{\lambda}} \cdot \vec{n}_i \,, \quad (25)$$

for an appropriate linearised advection velocity $\tilde{\vec{\lambda}}$, leads to

$$\oint_{\partial\triangle} u' \, \vec{\lambda} \cdot \vec{n} \; d\Gamma = \frac{1}{2}(u_i - u_j)\tilde{\vec{\lambda}} \cdot \vec{n}_j + \frac{1}{2}(u_i - u_k)\tilde{\vec{\lambda}} \cdot \vec{n}_k \quad (26)$$
$$+ \int_i^j \alpha_{ij}(u_i - u_j) \, \vec{\lambda} \cdot \vec{n} \; d\Gamma - \int_j^k \alpha_{jk}(u_i - u_j) \, \vec{\lambda} \cdot \vec{n} \; d\Gamma$$
$$- \int_k^i \alpha_{ki}(u_i - u_k) \, \vec{\lambda} \cdot \vec{n} \; d\Gamma + \int_j^k \alpha_{jk}(u_i - u_k) \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \,,$$

where i, j and k represent the three vertices of the (sub)cell under consideration and are chosen arbitrarily, but taken in an anticlockwise sense. This immediately gives

$$\oint_{\partial\triangle} u' \, \vec{\lambda} \cdot \vec{n} \; d\Gamma = \left[\frac{1}{2}\tilde{\vec{\lambda}} \cdot \vec{n}_j + \int_i^j \alpha_{ij} \, \vec{\lambda} \cdot \vec{n} \; d\Gamma - \int_j^k \alpha_{jk} \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \right] (u_i - u_j)$$
$$+ \left[\frac{1}{2}\tilde{\vec{\lambda}} \cdot \vec{n}_k - \int_k^i \alpha_{ki} \, \vec{\lambda} \cdot \vec{n} \; d\Gamma + \int_j^k \alpha_{jk} \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \right] (u_i - u_k)$$
$$= K_{ij}(u_i - u_j) + K_{ik}(u_i - u_k) \,. \quad (27)$$

Since $\alpha(\vec{x})$ is bounded, so are K_{ij} and K_{ik}. In fact, simple bounds can be derived, for example,

$$
\begin{aligned}
|K_{ij}| &= \left| \frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_j + \int_i^j \alpha \, \vec{\lambda} \cdot \vec{n} \; d\Gamma - \int_j^k \alpha \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \right| \\
&\leq \left| \frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_j \right| + \left| \int_i^j \alpha \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \right| + \left| \int_j^k \alpha \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \right| \\
&\leq \left| \frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_j \right| + \int_i^j |\alpha| \, |\vec{\lambda} \cdot \vec{n}| \; d\Gamma + \int_j^k |\alpha| \, |\vec{\lambda} \cdot \vec{n}| \; d\Gamma \\
&\leq \left| \frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_j \right| + C \int_i^j |\vec{\lambda} \cdot \vec{n}| \; d\Gamma + C \int_j^k |\vec{\lambda} \cdot \vec{n}| \; d\Gamma \\
&\leq \left| \frac{1}{2} \tilde{\vec{\lambda}} \cdot \vec{n}_j \right| + C \left[\max_{i \to j} |\vec{\lambda} \cdot \vec{n}_k| + \max_{j \to k} |\vec{\lambda} \cdot \vec{n}_i| \right] .
\end{aligned}
\tag{28}
$$

Similar expressions can be derived for K_{ik}. It then follows that the limited fluctuation, in the form (27), can be used to produce a locally positive update to the dependent variable u as long as Δt is small enough, and that the bound on the stable time-step satisfies $\Delta t = \mathcal{O}(C^{-1})$.

Remark: This proof is valid for interpolants of any order of accuracy, but for clarity the presentation below will be based on a third order (quadratic) interpolant. It is conceptually straightforward to extend this to higher order situations.

4.2 A Limited Third Order Scheme

Consider the two-dimensional scalar advection equation

$$
u_t + \vec{\lambda} \cdot \vec{\nabla} u = 0
\tag{29}
$$

where, for the purposes of this illustration, $\vec{\lambda}$ is assumed to vary linearly in space, u is assumed to vary in a piecewise quadratic manner in space, and $\vec{\nabla} \cdot \vec{\lambda} = 0$. This last condition means that the fluctuation can be written

$$
\phi = -\int_\triangle \vec{\lambda} \cdot \vec{\nabla} u \; d\Omega = -\int_\triangle \vec{\nabla} \cdot (\vec{\lambda} u) \; d\Omega = \oint_{\partial\triangle} u \, \vec{\lambda} \cdot \vec{n} \; d\Gamma ,
\tag{30}
$$

where $\vec{n}$ is the inward pointing normal. Since u is quadratic and $\vec{\lambda}$ is linear on the triangle, applying Simpson's rule along each edge of the triangle evaluates the fluctuation exactly, giving

$$\oint_{\partial\triangle} u\,\vec{\lambda}\cdot\vec{n}\,d\Gamma = \frac{1}{6}\left[u_i\vec{\lambda}_i\cdot\vec{n}_k + 4\left(\frac{u_i+u_j}{2}\right)\vec{\lambda}_{ij}\cdot\vec{n}_k + u_j\vec{\lambda}_j\cdot\vec{n}_k\right.$$
$$\left.+ 4\left(u_{ij} - \frac{u_i+u_j}{2}\right)\vec{\lambda}_{ij}\cdot\vec{n}_k\right]$$
$$+ \frac{1}{6}\left[u_j\vec{\lambda}_j\cdot\vec{n}_i + 4\left(\frac{u_j+u_k}{2}\right)\vec{\lambda}_{jk}\cdot\vec{n}_i + u_k\vec{\lambda}_k\cdot\vec{n}_i\right.$$
$$\left.+ 4\left(u_{jk} - \frac{u_j+u_k}{2}\right)\vec{\lambda}_{jk}\cdot\vec{n}_i\right]$$
$$+ \frac{1}{6}\left[u_k\vec{\lambda}_k\cdot\vec{n}_j + 4\left(\frac{u_k+u_i}{2}\right)\vec{\lambda}_{ki}\cdot\vec{n}_j + u_i\vec{\lambda}_i\cdot\vec{n}_j\right.$$
$$\left.+ 4\left(u_{ki} - \frac{u_k+u_i}{2}\right)\vec{\lambda}_{ki}\cdot\vec{n}_j\right],\qquad(31)$$

where i, j and k are again the vertices of the triangular (sub)cell, taken anticlockwise. The values u_{ij}, u_{jk} and u_{ki} represent the interpolated values of the dependent variable at the midpoints of each of the cell (for the Caraeni scheme) or subcell (for the Abgrall-Roe scheme) edges. For higher order representations of the interpolant the above expressions become more complicated due to the terms required for the additional quadrature points used to integrate the fluctuation exactly.

At this point it is still possible to limit the (sub)cell fluctuation as a whole rather than considering the interpolant along each edge separately, as will be done here. Although it would most likely lead to a discontinuous representation of the dependent variable, this may not be a problem in practice, as long as all of these discontinuities lie in regions where the solution itself is smooth [5]. The results illustrated in Figure 16 suggest that this may indeed be the case but this possibility will be left as future work because the analysis of Section 4.1 suggests that the fluctuation should be split in to edge-based components, as is done in the N scheme, and as soon as an edge-based limiting procedure is considered, the continuous representation used here results automatically.

Now note that, for **linear** u and linear $\vec{\lambda}$ satisfying $\vec{\nabla}\cdot\vec{\lambda}=0$, the fluctuation can be written

$$\oint_{\partial\triangle} u\,\vec{\lambda}\cdot\vec{n}\,d\Gamma = \frac{1}{6}\left[u_i\vec{\lambda}_i\cdot\vec{n}_k + 4\left(\frac{u_i+u_j}{2}\right)\vec{\lambda}_{ij}\cdot\vec{n}_k + u_j\vec{\lambda}_j\cdot\vec{n}_k\right]$$
$$+ \frac{1}{6}\left[u_j\vec{\lambda}_j\cdot\vec{n}_i + 4\left(\frac{u_j+u_k}{2}\right)\vec{\lambda}_{jk}\cdot\vec{n}_i + u_k\vec{\lambda}_k\cdot\vec{n}_i\right]$$
$$+ \frac{1}{6}\left[u_k\vec{\lambda}_k\cdot\vec{n}_j + 4\left(\frac{u_k+u_i}{2}\right)\vec{\lambda}_{ki}\cdot\vec{n}_j + u_i\vec{\lambda}_i\cdot\vec{n}_j\right]\qquad(32)$$

so, since in this case

$$\int_\triangle \vec{\lambda}\cdot\vec{\nabla}u\,d\Omega = \left(\int_\triangle \vec{\lambda}\,d\Omega\right)\cdot\vec{\nabla}u = S_\triangle\tilde{\vec{\lambda}}\cdot\vec{\nabla}u,\qquad(33)$$

where $\tilde{\vec{\lambda}} = (\vec{\lambda}_i + \vec{\lambda}_j + \vec{\lambda}_k)/3$, it follows from (3) that in the high order case

$$
\oint_{\partial \triangle} u\, \vec{\lambda} \cdot \vec{n} \; d\Gamma = -\frac{1}{2} u_i \tilde{\vec{\lambda}} \cdot \vec{n}_i - \frac{1}{2} u_j \tilde{\vec{\lambda}} \cdot \vec{n}_j - \frac{1}{2} u_k \tilde{\vec{\lambda}} \cdot \vec{n}_k
$$

$$
+ \frac{2}{3} \left(u_{ij} - \frac{u_i + u_j}{2} \right) \vec{\lambda}_{ij} \cdot \vec{n}_k \; + \; \frac{2}{3} \left(u_{jk} - \frac{u_j + u_k}{2} \right) \vec{\lambda}_{jk} \cdot \vec{n}_i
$$

$$
+ \frac{2}{3} \left(u_{ki} - \frac{u_k + u_i}{2} \right) \vec{\lambda}_{ki} \cdot \vec{n}_j \, . \tag{34}
$$

As it stands this fluctuation, evaluated over a mesh cell for the Caraeni approach or a subcell for the Abgrall-Roe scheme, cannot be distributed to the vertices denoted by the indices i, j and k in a manner guaranteed to be positive. In order to allow this, the interpolated values at the midpoints of each of the (sub)cell edges, u_{ij}, u_{jk} and u_{ki} are limited so that they satisfy (20) and (22). For cubic or higher order representations the limiting is applied to the values at all of the quadrature points along each (sub)cell edge, *i.e.* the quadrature points on the subcell edges for the Abgrall-Roe scheme but on the full cell edge for the Caraeni approach.

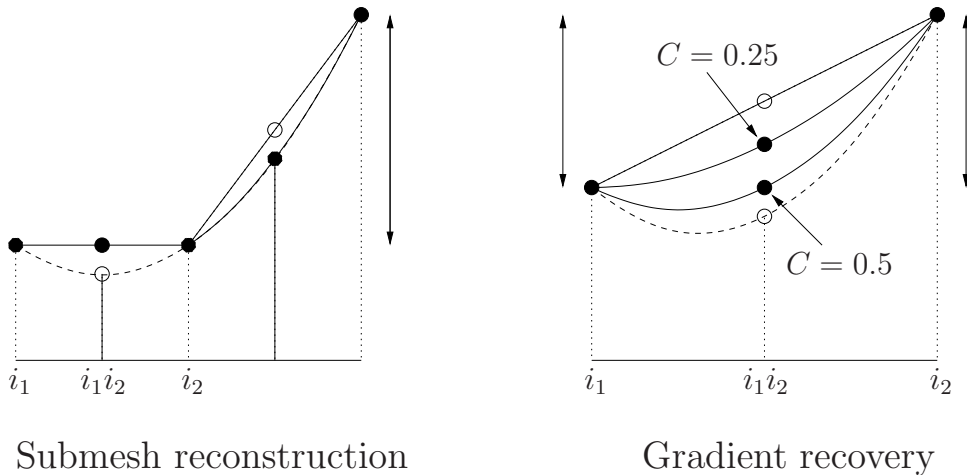Submesh reconstruction   Gradient recovery

Fig. 8. The effect of limiting the reconstruction along a cell edge. The dotted lines are the linear interpolants $\bar{u}(\vec{x})$, the dashed lines the quadratic interpolants $u(\vec{x})$ and the solid lines the limited interpolants $u'(\vec{x})$ for $C = 0.25$ and $C = 0.5$ (they are the same in the left hand diagram). The double headed arrows indicate the range of values that the polynomial is allowed to take at the quadrature point and the filled circles show the *limited* values at the quadrature points.

The midpoint values are limited so that new values, u'_{ij}, u'_{jk} and u'_{ki}, are created for the dependent variable which, following (20), satisfy

$$\left| u'_{ij} - \frac{u_i + u_j}{2} \right| \le C \left| u_i - u_j \right|$$

$$\left| u'_{jk} - \frac{u_j + u_k}{2} \right| \le C \left| u_j - u_k \right|$$

$$\left| u'_{ki} - \frac{u_k + u_i}{2} \right| \le C \left| u_k - u_i \right|, \tag{35}$$

where $C \ge 0$ is some specified constant. In the general case, the higher order polynomial would be limited to ensure that the above relations were satisfied at every quadrature point, not just the midpoints. The optimal choice for C remains an open question, but three significant values are

- $C = 0$, which simply returns the linear representation (and, ultimately in the proposed scheme, leads to the PSI).
- $C = 0.25$, which is the largest value that guarantees that the limited interpolant along each edge is monotonic (in the quadratic case).

Monotonicity can be proved simply by considering the Newton interpolant of the three data points (x_L, u_L), (x_M, u_M) and (x_R, u_R), where

$$x_M = \frac{x_L + x_R}{2} \qquad \text{and} \qquad u_M = \frac{u_L + u_R}{2} + C(u_R - u_L) \tag{36}$$

for some constant C. These represent the quadrature points at which the interpolant is evaluated along the edge of the triangle when Simpson's rule is used. They are considered in one dimension for simplicity. Constructing the Newton interpolant leads to the equation

$$u = u_L + (2C + 1)\frac{\Delta u}{\Delta x}(x - x_L) - \frac{4C}{\Delta x}\frac{\Delta u}{\Delta x}(x - x_L)(x - x_M) \tag{37}$$

in which $\Delta u = u_R - u_L$ and $\Delta x = x_R - x_L$, and it immediately follows that

$$\frac{\mathrm{d}u}{\mathrm{d}x} = (2C + 1)\frac{\Delta u}{\Delta x} - \frac{4C}{\Delta x}\frac{\Delta u}{\Delta x}(2x - x_L - x_M). \tag{38}$$

Now, for the interpolant to be monotonic, it is necessary that $\frac{\mathrm{d}u}{\mathrm{d}x}$ always has the same sign as $\frac{\Delta u}{\Delta x}$, which leads, after some algebraic manipulation, to the inequality

$$\frac{4C}{\Delta x}\left((x - x_L) + (x - x_M) - \frac{\Delta x}{2} \right) \le 1. \tag{39}$$

This is a linear equation in x, so it suffices to impose (39) at the endpoints of the domain, x_L and x_R, to find the admissible range of C. It then follows immediately that $0 \le |C| \le 0.25$, *i.e.* $C \le 0.25$ in (35). The geometric effect of using $C = 0.25$ is illustrated in Figure 8.

This is the value which is used in the rest of this work. Using larger values of C tends to reduce the rate of convergence to the steady state, even when the same time-step is used for the calculations.

- $C = 0.5$, which is the largest value that guarantees that the limited mid-point values are bounded by the endpoint values for any given edge (in the quadratic case).

This follows immediately from imposing

$$\min(u_L, u_R) \ \leq \ u_M \ = \ \frac{u_L + u_R}{2} + C(u_R - u_L) \ \leq \ \max(u_L, u_R) \qquad (40)$$

and the geometric effect of using this value is illustrated in Figure 8.

Note though that the time-step restriction depends on C, *cf.* Equation (28), and that it becomes more severe as C increases, *i.e.* $\Delta t \to 0$ as $C \to \infty$.

In this work, the limiting is simply carried out on an edge $i_1 i_2$ by setting

$$u'_{i_1 i_2} \ = \ \frac{u_{i_1} + u_{i_2}}{2} + \alpha_{i_1 i_2}(u_{i_1} - u_{i_2}) \qquad (41)$$

where

$$\alpha_{i_1 i_2} \ = \ \max\left(-C, \min\left[C, \frac{u_{i_1 i_2} - (u_{i_1} + u_{i_2})/2}{u_{i_1} - u_{i_2}}\right]\right), \qquad (42)$$

in which division by zero is carefully avoided. For a single quadrature point this procedure is equivalent to setting $\delta u'_{i_1 i_2}(\vec{x}) = C'_{i_1 i_2} \delta u_{i_1 i_2}(\vec{x})$ where (42) ensures that $C'_{i_1 i_2} \in [0, 1]$ for each edge. Once more, in the general case this constraint would be imposed at each of the quadrature points and then, for example, the minimum value of α could be taken on each edge.

If a discontinuous representation was considered, allowing the limiting to be carried out cell-by-cell instead of edge-by-edge, then the interpolant at each quadrature point would still take a similar form to (41), but α might need to be modified, to account for any changes in the constraints used to impose local positivity. However, given that the positivity analysis typically decomposes the fluctuation in to edge-based components, the continuous representation is actually quite natural. In either situation, though, it is clear that the limiting of the interpolant will tend to be applied in regions where the solution is varying rapidly. In particular, it will tend to return a low order scheme close to discontinuities, as is also expected when the more traditional one-dimensional limiters are used.

The modification given by (41) and (42) means that the fluctuation now takes the form

$$\oint_{\partial\triangle} u' \, \vec{\lambda} \cdot \vec{n} \, d\Gamma = -\frac{1}{2} u_i \tilde{\vec{\lambda}} \cdot \vec{n}_i - \frac{1}{2} u_j \tilde{\vec{\lambda}} \cdot \vec{n}_j - \frac{1}{2} u_k \tilde{\vec{\lambda}} \cdot \vec{n}_k \qquad (43)$$

$$+ \frac{2}{3}\alpha_{ij}(u_i - u_j)\vec{\lambda}_{ij} \cdot \vec{n}_k \ + \ \frac{2}{3}\alpha_{jk}(u_j - u_k)\vec{\lambda}_{jk} \cdot \vec{n}_i \ + \ \frac{2}{3}\alpha_{ki}(u_k - u_i)\vec{\lambda}_{ki} \cdot \vec{n}_j \, ,$$

22

which, according to the analysis of Section 4.1, can be distributed to the vertices i, j and k, in a positive manner. It remains to determine the best way of doing this.

It is worth noting here that, when a piecewise linear representation of the dependent variable is used, the fluctuation can always (even in the single target case) be written as

$$-\frac{1}{2}u_i\tilde{\vec{\lambda}}\cdot\vec{n}_i - \frac{1}{2}u_j\tilde{\vec{\lambda}}\cdot\vec{n}_j - \frac{1}{2}u_k\tilde{\vec{\lambda}}\cdot\vec{n}_k = \frac{1}{2}(u_i - u_j)\tilde{\vec{\lambda}}\cdot\vec{n}_j + \frac{1}{2}(u_i - u_k)\tilde{\vec{\lambda}}\cdot\vec{n}_k$$
$$= \phi_{ij} + \phi_{ik}\,, \tag{44}$$

where the vertices i, j and k are chosen so that the inflow parameters $k_j = \frac{1}{2}\tilde{\vec{\lambda}}\cdot\vec{n}_j$ and $k_k = \frac{1}{2}\tilde{\vec{\lambda}}\cdot\vec{n}_k$ have the same sign (or are zero).

Remark: In the special case where the flow is parallel to an edge of the (sub)cell one of k_j or k_k will be zero and the split illustrated by Equation (44) is not unique. In the piecewise linear (second order) case, this coincides with the situation where one of the edge contributions, ϕ_{ij} or ϕ_{ik} is zero, so even though the vertices represented by i, j and k switch as the advection velocity rotates through this orientation, when the upwind distribution of, for example, the N scheme is applied, the distribution still depends continuously on the advection velocity. The situation is more complicated when a higher order representation is used.

It follows immediately from (44) and the definition of the inflow parameters below it that the low order fluctuation is given by

$$\phi^{LO} = k_j(u_i - u_j) + k_k(u_i - u_k) \tag{45}$$

and that the N scheme (in both its one-target and two-target incarnations) can be viewed as distributing ϕ_{ij} via

$$S_i u_i \;\rightarrow\; S_i u_i + \frac{1}{2}\Delta t\,\tilde{\vec{\lambda}}\cdot\vec{n}_j(u_i - u_j) \quad \text{if} \quad \tilde{\vec{\lambda}}\cdot\vec{n}_j \le 0$$
$$S_j u_j \;\rightarrow\; S_j u_j + \frac{1}{2}\Delta t\,\tilde{\vec{\lambda}}\cdot\vec{n}_j(u_i - u_j) \quad \text{if} \quad \tilde{\vec{\lambda}}\cdot\vec{n}_j > 0\,. \tag{46}$$

A similar approach can be used to distribute ϕ_{ik}.

Since the fluctuation can be written

$$\phi = k_j^+(u_i - u_j) + k_j^-(u_i - u_j) + k_k^+(u_i - u_k) + k_k^-(u_i - u_k)\,, \tag{47}$$

where $[\;]^+$ and $[\;]^-$ are, respectively, the positive and negative parts of the

argument within the square brackets, an equivalent form of this distribution is given by

$$
\begin{aligned}
S_i u_i &\rightarrow S_i u_i \;+\; \Delta t \left[k_j^-\left(u_i - u_j\right) + k_k^-\left(u_i - u_k\right)\right] \\
S_j u_j &\rightarrow S_j u_j \;+\; \Delta t \, k_j^+\left(u_i - u_j\right) \\
S_k u_k &\rightarrow S_k u_k \;+\; \Delta t \, k_k^+\left(u_i - u_k\right).
\end{aligned}
\tag{48}
$$

In the limited high order case

$$
\begin{aligned}
\oint_{\partial\triangle} u' \, \vec{\lambda} \cdot \vec{n} \; d\Gamma = {} & \frac{1}{2}(u_i - u_j)\tilde{\vec{\lambda}} \cdot \vec{n}_j \;+\; \frac{2}{3}\alpha_{ij}(u_i - u_j)\vec{\lambda}_{ij} \cdot \vec{n}_k \\
& + \frac{1}{2}(u_i - u_k)\tilde{\vec{\lambda}} \cdot \vec{n}_k \;+\; \frac{2}{3}\alpha_{ki}(u_k - u_i)\vec{\lambda}_{ki} \cdot \vec{n}_j \\
& \hspace{4cm} + \frac{2}{3}\alpha_{jk}(u_j - u_k)\vec{\lambda}_{jk} \cdot \vec{n}_i
\end{aligned}
\tag{49}
$$

so, since

$$
u_j - u_k \;\equiv\; (u_j - u_i) + (u_i - u_k) \;\equiv\; (u_i - u_k) - (u_i - u_j)
\tag{50}
$$

it follows that the limited fluctuation, denoted here by the superscript LIM, can be written

$$
\begin{aligned}
\phi^{LIM} = {} & \oint_{\partial\triangle} u' \, \vec{\lambda} \cdot \vec{n} \; d\Gamma \\
= {} & \frac{1}{2}(u_i - u_j)\tilde{\vec{\lambda}} \cdot \vec{n}_j \;+\; \frac{2}{3}\alpha_{ij}(u_i - u_j)\vec{\lambda}_{ij} \cdot \vec{n}_k \;-\; \frac{2}{3}\alpha_{jk}(u_i - u_j)\vec{\lambda}_{jk} \cdot \vec{n}_i \\
& + \frac{1}{2}(u_i - u_k)\tilde{\vec{\lambda}} \cdot \vec{n}_k \;-\; \frac{2}{3}\alpha_{ki}(u_i - u_k)\vec{\lambda}_{ki} \cdot \vec{n}_j \;+\; \frac{2}{3}\alpha_{jk}(u_i - u_k)\vec{\lambda}_{jk} \cdot \vec{n}_i \\
= {} & K_{ij}(u_i - u_j) \;+\; K_{ik}(u_i - u_k)
\end{aligned}
\tag{51}
$$

where K_{ij} and K_{ik} can be found easily, *cf.* (27). This now has a similar form to the linear fluctuation (45).

If K_{ij} has the same sign as k_j in (45), sending $K_{ij}(u_i - u_j)$ to the same node as $k_j(u_i - u_j)$ clearly leads to a locally positive distribution. If K_{ij} and k_j have opposite signs then the distribution can be reversed, *i.e.* $K_{ij}(u_i - u_j)$ can be sent to the opposite node on edge ij to ensure local positivity. The fluctuation associated with edge ik can be treated in a similar manner.

Remark: When $k_j = 0$ (or $k_k = 0$) the N scheme gives no guidance on the direction in which the fluctuation associated with edge ij (or ik) should be sent. In fact, due to the earlier limiting of the interpolant (35), it doesn't matter

how the vertices i and j (or k) are chosen in (45), from the point of view of ensuring positivity (in practice, it is assumed that the N scheme would have sent its fluctuation in the direction suggested by $k_j > 0$ (or $k_k > 0$), and the procedure described in the previous paragraph is followed). However, this is the situation highlighted by the previous remark, in which the node to which the edge contribution is sent switches: for the N scheme it coincides with that contribution being zero, so the distribution varies continuously as the advection velocity rotates, but for the high order case this is no longer the case. The result is that the new scheme does not depend continuously on the advection velocity, though this doesn't appear to have any detrimental effect on the numerical results, particularly the rate of convergence to the steady state (see Section 5). Alternative, continuous, distributions have been considered, but the results were invariably far less accurate than those obtained using this approach.

Since K_{ij} and K_{ik} are bounded, the above approach automatically leads to a positive scheme for small enough Δt. Following the formulation of the N scheme given by (45)–(48), the contributions due to edge ij,

$$
\begin{aligned}
S_i u_i \;&\rightarrow\; S_i u_i + \Delta t\, K_{ij}(u_i - u_j) \quad && \text{if} \quad && K_{ij} \le 0 \\
S_j u_j \;&\rightarrow\; S_j u_j + \Delta t\, K_{ij}(u_i - u_j) \quad && \text{if} \quad && K_{ij} > 0
\end{aligned}
\tag{52}
$$

clearly lead to a positive distribution for small enough Δt and, in the general case, the fluctuation can be written

$$
\phi \;=\; K_{ij}^{+}(u_i - u_j) \;+\; K_{ij}^{-}(u_i - u_j) \;+\; K_{ik}^{+}(u_i - u_k) \;+\; K_{ik}^{-}(u_i - u_k). \tag{53}
$$

The distribution therefore takes the form

$$
\begin{aligned}
S_i u_i &\rightarrow S_i u_i \;+\; \Delta t\left[K_{ij}^{-}(u_i - u_j) + K_{ik}^{-}(u_i - u_k) \right] \\
S_j u_j &\rightarrow S_j u_j \;+\; \Delta t\, K_{ij}^{+}(u_i - u_j) \\
S_k u_k &\rightarrow S_k u_k \;+\; \Delta t\, K_{ik}^{+}(u_i - u_k),
\end{aligned}
\tag{54}
$$

where the vertices i, j and k have been chosen according to the inflow edges dictated by the N scheme. This is clearly locally positive for

$$
\Delta t \;\le\; \min\left(\frac{-S_i}{K_{ij}^{-} + K_{ik}^{-}},\, \frac{S_j}{K_{ij}^{+}},\, \frac{S_k}{K_{ik}^{+}} \right), \tag{55}
$$

which implies global positivity, under a slightly different constraint on the time-step. It is also worth recalling that *any* fluctuation which can be written in the form

$$
\phi \;=\; K_{ij}'(u_i - u_j) \;+\; K_{jk}'(u_j - u_k) \;+\; K_{ki}'(u_k - u_i) \tag{56}
$$

where each of the K' are bounded, can be distributed in a positive manner. There are clearly many possible alternatives to the method described above but only this one, which has so far turned out to be the most successful, will be considered here.

The form of the fluctuation given in (56) suggests that the scheme is compact, but it should be remembered that this is only the case for the schemes based on a linear representation of the data, for which the K' depend only on the local data u_i, u_j and u_k. All of the higher order schemes presented gather information from further away so the scheme is no longer local, although in the submesh reconstruction the stencil used to construct the fluctuation does remain local to the mesh cell (not the subcell).

Once these contributions for the two edges have been gathered together to give contributions from (sub)cell j to its vertices, the distribution coefficients of the resulting N-like scheme, denoted by the superscript N^*, take the form

$$(\phi_i^j)^{N^*} \;=\; (\alpha_i^j)^{N^*}\, \phi_j^{LIM}\,. \tag{57}$$

These can be limited in precisely the manner which created the PSI scheme (11) by imposing linearity preservation on the N scheme, *i.e.*

$$(\phi_i^j)^{PSI^*} \;=\; \frac{[(\alpha_i^j)^{N^*}]^+}{\sum_{k\in\triangle_j}[(\alpha_k^j)^{N^*}]^+}\, \phi_j^{LIM} \;=\; (\alpha_i^j)^{PSI^*}\, \phi_j^{LIM}\,. \tag{58}$$

The form of the distribution given in (57) ensures that at least one distribution coefficient within each (sub)cell must be positive so, unlike with the Abgrall-Roe scheme, conservation is assured without the need for any modifications. As with the PSI scheme, the limiting procedure will never increase the magnitude of the distribution coefficients, so the positivity condition (55) is actually stronger than necessary.

Results: The results of applying the limiting to the two high order methods are shown in Figures 9 and 10, using the same meshes as were used before to obtain results for their respective basic schemes, described in Sections 3.1 and 3.2. The oscillations have clearly been removed in each case (the contour plots show the smoothness of the solutions and the minimum and maximum solution values are recreated exactly, to machine precision), at the expense of a drop in peak value in the smooth test case. A more detailed comparison with the other schemes is provided in the following section.
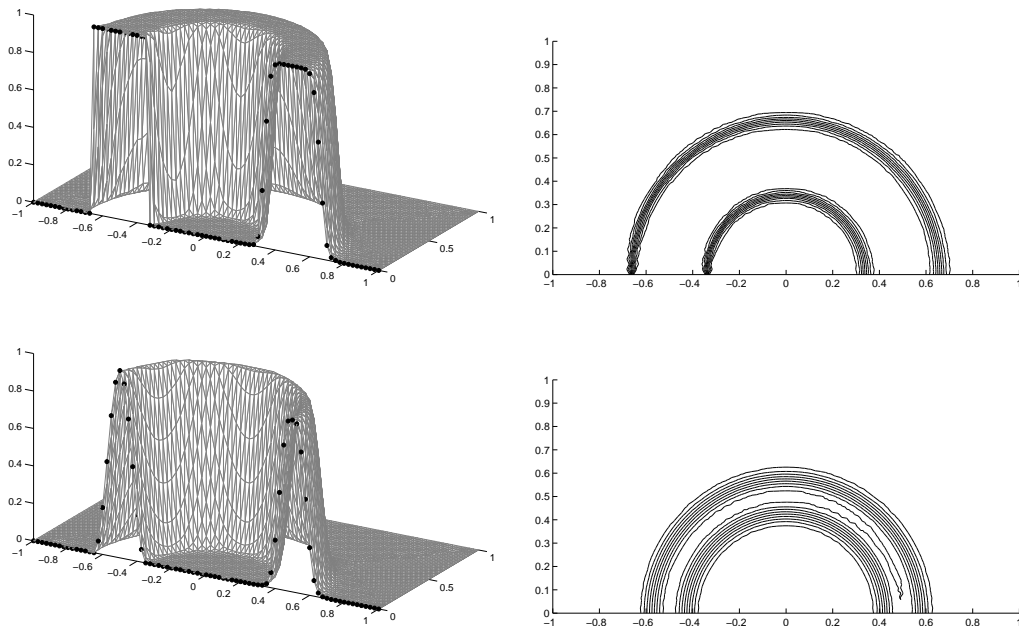
Fig. 9. The modified Abgrall-Roe scheme applied to Test Case A (top) and Test Case B (bottom).
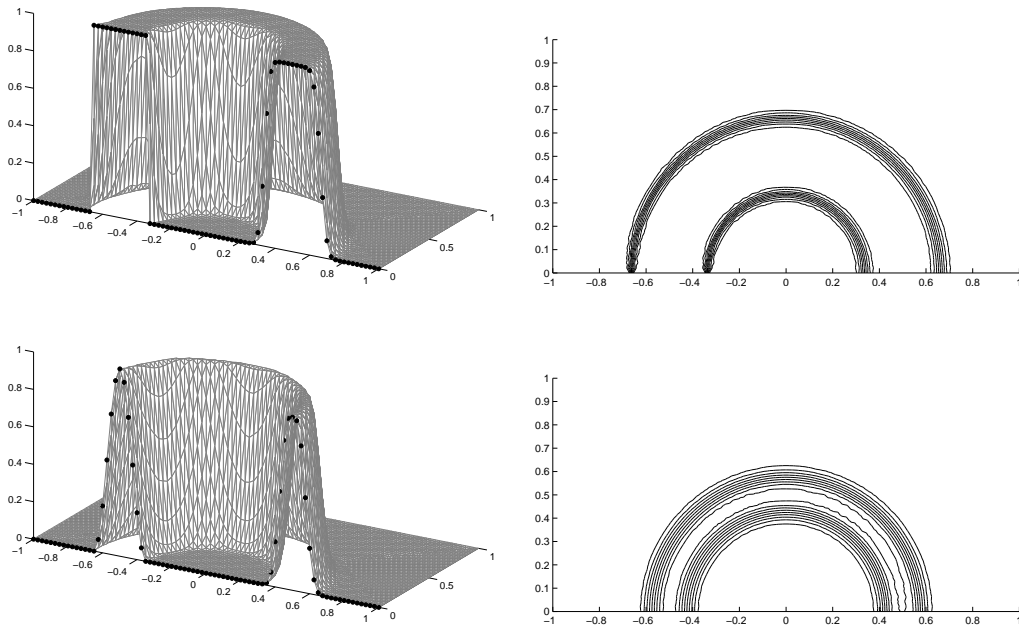

Fig. 10. The modified Caraeni scheme applied to Test Case A (top) and Test Case B (bottom).

5 Advection Results Comparison

Results have been shown in earlier sections to illustrate the behaviour of the methods described. Here, they will be compared, not only for test cases A and B (defined in Section 2.1) but also for a third situation (Test Case C), which uses exactly the same velocity field but a smoother solution profile, given by

$$u(x,y,0) = \begin{cases} G(x) & \text{for } -0.75 \le x \le -0.25, \ y = 0, \\ 0 & \text{otherwise} \end{cases} \tag{59}$$

in which

$$G(x) = \begin{cases} g(4x+3) & \text{for } -0.75 \le x \le -0.5 \\ g(-4x-1) & \text{for } -0.5 < x \le -0.25 \end{cases} \tag{60}$$

where

$$g(x) = x^5(70x^4 - 315x^3 + 540x^2 - 420x + 126). \tag{61}$$

The exact solution to this problem, $u(x,y) = G(r)$ for $0.25 \le r \le 0.75$ with $u(x,y) = 0$ elsewhere, has continuous fourth derivative so it is used to test order of accuracy in the presence of turning points in the solution, using a non-constant (in space) advection velocity on both structured and genuinely unstructured (but uniform) triangular meshes.

The results obtained for the different methods are compared in Tables 1 and 2 and Figures 11–14. For the purposes of this comparison the new approach will be known as the Submesh PSI scheme when it is applied to the Abgrall-Roe scheme and as the Gradient PSI scheme when it is applied to Caraeni's method.

It is immediately apparent that the unphysical oscillations have been removed completely in each case. There is very little difference between the quality of the profile at outflow for the two schemes (see Figures 11 and 12), though both are significantly better than the PSI scheme. In the smooth test case (B), Caraeni's scheme is clearly the best, but this is at the expense of the oscillations that are visible close to discontinuities in test case A, which do not reduce in amplitude as the mesh is refined.

The results shown in Figure 13 and Table 1 were obtained using a series of uniform *unstructured* meshes, obtained using the same mesh generator, but successively halving the background mesh size specified. This illustrates the rate at which the error would decrease as finer meshes are used which are not embedded in the coarser ones, a situation which is commonly encountered when seeking more accurate solutions. The errors obtained on the finest pair of meshes were used to estimate the order of accuracy obtained and shown in the final two columns of Table 1. The same comparison was carried out on

28

a series of structured triangular meshes (a square mesh was subdivided in a manner whereby the diagonals alternated in direction from cell to cell) and the corresponding results are shown in Figure 14 and Table 2.

On the unstructured meshes it is clear that all of the high order methods lose their formal order of accuracy, particularly those which use a submesh reconstruction. This is not surprising because the connectivity of the mesh changes as the refinements are carried out, leading to a "bumpiness" in the reduction of the errors and unreliable estimates of the orders of accuracy in Table 1. In Caraeni's scheme, for example, the approach used to reconstruct the nodal solution gradients, given by Equations (16) and (17), drops in accuracy so third order is no longer expected. Even so, the schemes based on gradient recovery are significantly more accurate on the meshes used than any of the others. On very fine meshes the submesh reconstruction schemes actually give higher L_∞ errors than the PSI scheme. For the original Abgrall-Roe scheme this seems to be because it is reverting to the unstable central distribution at key nodes, whereas the Submesh PSI scheme may lose accuracy because the interpolant is being limited separately on each of the four subcells, affecting the smoothness of the interpolant.

Higher orders of accuracy are obtained on the structured meshes, though the theoretical values are not always achieved. There is, for example, still a loss of accuracy in the approximation to the solution gradients at the boundary nodes, which reduces the order in the L_∞ norm for the gradient reconstruction schemes. Also, the effect of the occasional use of a central difference is still visible in the L_∞ errors of the Abgrall-Roe scheme. The errors measured in the L_1 norm decrease at rates much closer to those predicted by the theory.

Overall, the gradient recovery methods still give more accurate results than those which use the submesh reconstruction. In addition, Caraeni's original scheme is the most accurate of all, which is expected because it uses an LDA distribution step, known to produce less numerical dissipation than the PSI scheme in the piecewise linear case. For a smooth solution, where high accuracy is more important than removing the possibility of small oscillations, such LDA-type schemes are ideal. Approaches based on the PSI scheme are more appropriate when discontinuities may occur. The very high order of accuracy seen in Table 2 for the L_1 error of the Gradient PSI scheme (higher even than the original Caraeni scheme) is anomalous and suggests that even on these fine meshes, convergence has not been reached. The limited scheme is always the less accurate of the two, even at the finest mesh resolution.

Although the imposition of positivity tends to reduce accuracy when applied to Caraeni's scheme (as expected) it improves the accuracy when applied to the method of Abgrall and Roe. This is most likely because, in its original form, the latter approach resorts to a central distribution in some circumstances which,

if used on its own, would be unstable. In contrast, the Gradient PSI scheme has been modified from an LDA distribution of an unlimited fluctuation. A small drop in accuracy is therefore expected in test cases such as this where the advection velocity is not constant in space and the solution profile has a turning point, where the scheme will typically return to the standard, second order, PSI scheme.

Table 1
Accuracy measures for the uniform unstructured triangular meshes (none of the "non-oscillatory" results goes negative). The figures shown in the first three columns have been obtained on the 3806 node and subdivided 984 node meshes shown in Figure 2.

	Test Case A		Test Case B	Test Case C	
Scheme	$\min(u)$	$\max(u)$	$\max(u)$ outflow	L_1 order	L_∞ order
N	0.0000	1.0000	0.5877	0.94	0.84
PSI	0.0000	1.0000	0.8355	1.93	1.89
Abgrall-Roe	-0.0593	1.0131	0.9702	1.46	1.01
Submesh PSI	0.0000	1.0000	0.9180	1.09	0.53
Caraeni	-0.1464	1.1301	0.9996	2.05	1.57
Gradient PSI	0.0000	1.0000	0.9242	1.83	2.10

Table 2
Accuracy measures for the uniform structured triangular meshes (none of the "non-oscillatory" results goes negative). The figures shown in the first three columns have been obtained on regular 65×33 node and subdivided 33×17 node triangular meshes.

	Test Case A		Test Case B	Test Case C	
Scheme	$\min(u)$	$\max(u)$	$\max(u)$ outflow	L_1 order	L_∞ order
N	0.0000	1.0000	0.4670	0.91	0.75
PSI	0.0000	1.0000	0.7751	1.92	1.81
Abgrall-Roe	-0.0951	1.0107	0.9143	2.57	1.54
Submesh PSI	0.0000	1.0000	0.8641	2.68	2.30
Caraeni	-0.2382	1.2036	1.0001	3.05	2.01
Gradient PSI	0.0000	1.0000	0.8832	3.47	1.23

The rates of convergence to the steady state of all of the calculations shown in Figures 3, 4, 6, 7, 9 and 10, along with those of the corresponding calculations for Test Case C, are illustrated in Figure 15. No convergence acceleration techniques, such as local time-stepping, have been applied. All of the methods except that of Abgrall and Roe converge rapidly to machine accuracy. Among the converging schemes, the N scheme is the most rapid in each case, and the
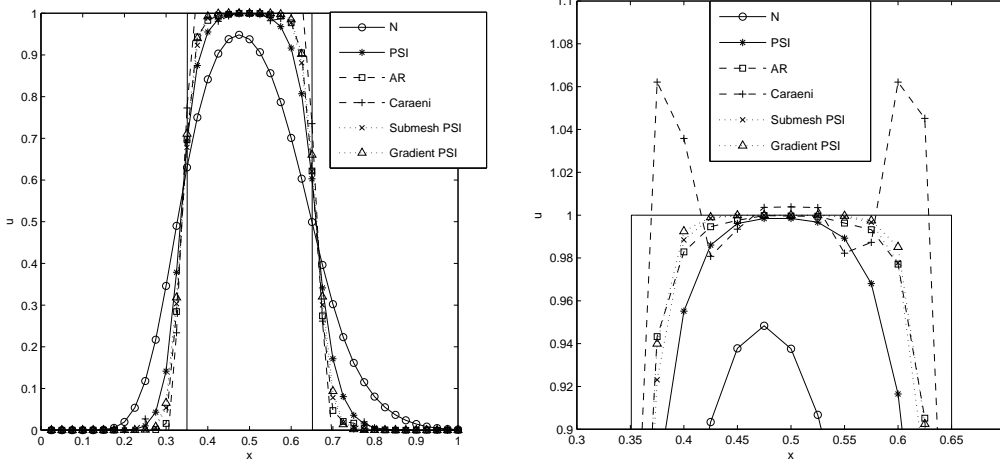
Fig. 11. Square wave profile at outflow (left) and magnified (right). The solid line without symbols is the exact solution.
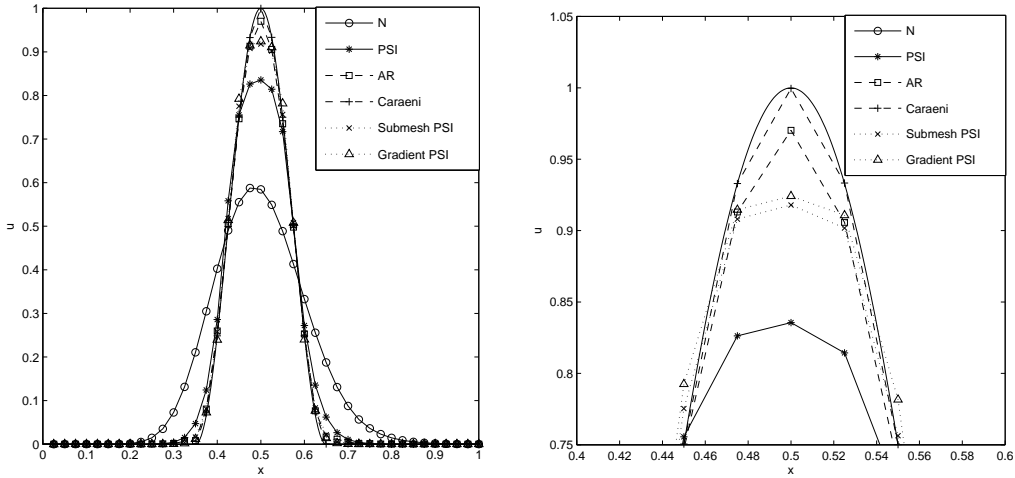

Fig. 12. Cosine-squared profile at outflow (left) and magnified (right). The solid line without symbols is the exact solution.

PSI scheme gives faster convergence when applied to the linear representation rather than any of the higher order representations. Otherwise, the identity of the most rapidly converging scheme depends on the level to which the residuals are required to be reduced. It should be noted that the results presented for the high order limited schemes are for $C = 0.25$. The simulations still converge to the same level for $C = 0.5$ but it invariably takes more than twice as long to do so, even when the same time-step is used, and provides little difference in the quality of the converged results. For higher values of C, the method ceases to converge to machine accuracy.

This rapid convergence is encouraging, since recent work by Abgrall [3] has shown that methods based on the nonlinear mappings, discussed in [8] and used in this paper (see Equations (11) and (58)), to impose linearity preservation on positive schemes might lead to attempts to solve ill-posed algebraic
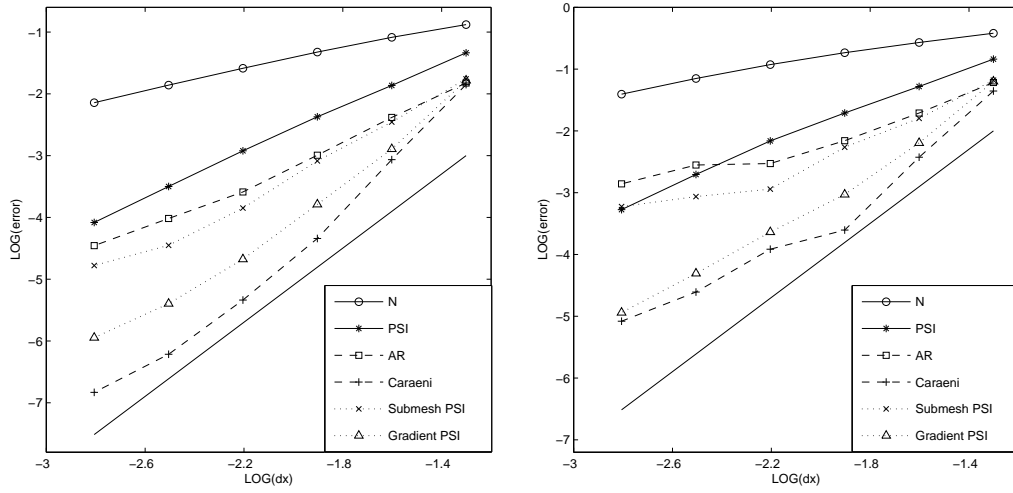
Fig. 13. Smooth polynomial profile approximated on the uniform unstructured triangular meshes: L_1 errors (left) and L_∞ errors (right). The solid line without symbols is of slope 3 in each case.
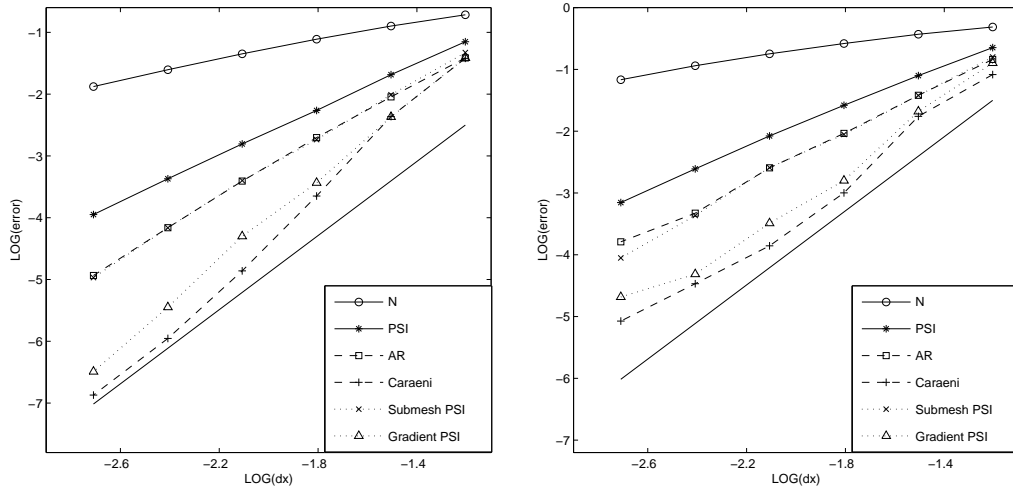

Fig. 14. Smooth polynomial profile approximated on the uniform structured triangular meshes: L_1 errors (left) and L_∞ errors (right). The solid line without symbols is of slope 3 in each case.

systems. The convergence of each experiment to machine accuracy and the relatively smooth contours seen in Figures 9 and 10 suggests that the algebraic systems, solved here by a pseudo-time iteration, are well posed. It must be emphasised though that only the scalar problem has been considered. It has so far proved difficult to maintain this rapid convergence when the PSI scheme has been extended to nonlinear systems of equations [3] and this is likely to remain true of the high order PSI schemes proposed here until a better generalisation can be found. In such situations schemes based on the LDA scheme gain a clear advantage in the simulation of smooth solutions [9].

It was mentioned at the end of Section 3 that a discontinuous representation of
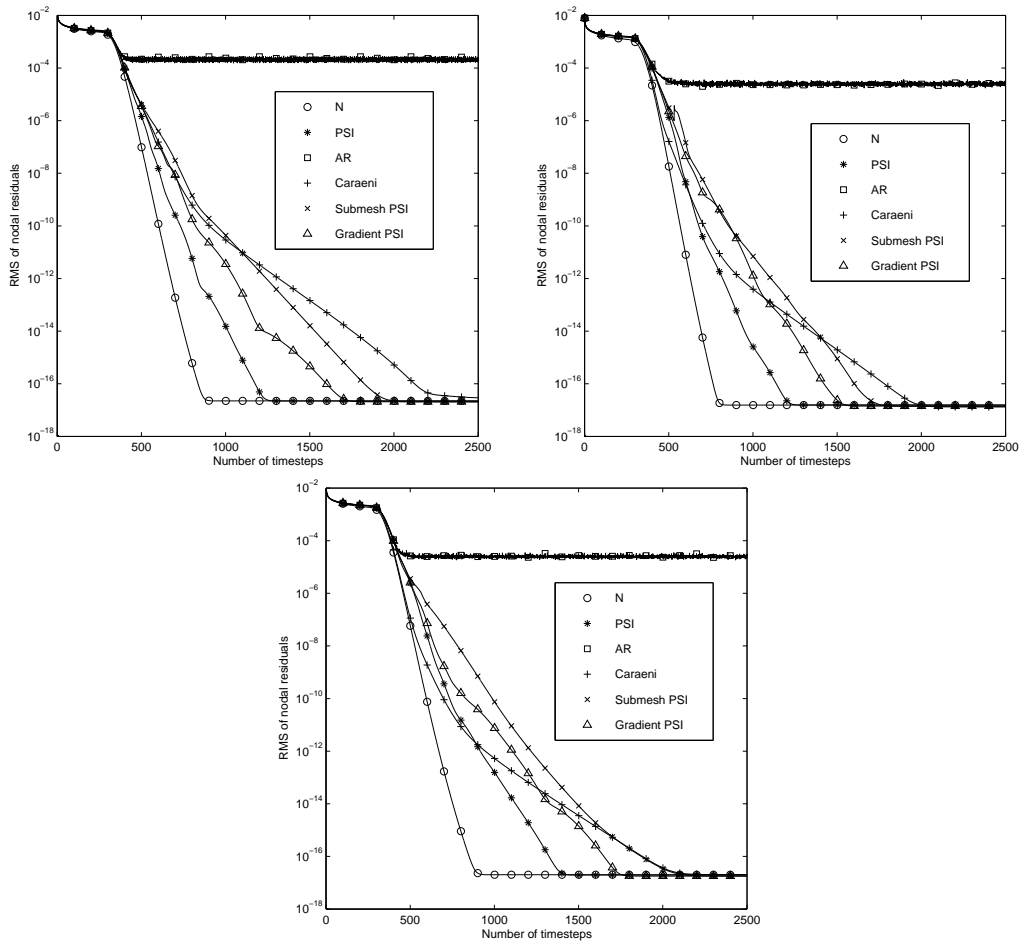
Fig. 15. Convergence histories for each of the methods on the meshes shown in Figure 2 for Test Cases A (top left), B (top right) and C (bottom).

the dependent variable could be used, as long as the discontinuities remained in regions where the underlying flow was smooth. It is therefore of interest to note where the limiting procedure affects the polynomial. Figure 16 provides an illustration of this by highlighting the edges along which the solution value at the quadrature point has been modified (and hence the representation has lost the full order of accuracy) for Test Case B. It shows that the affected edges are typically in regions where the solution is close to turning points, at peaks and close to zero on either side. It also shows that for this velocity profile and for a linear equation the region expands as the profile diffuses.

6 A Simple Nonlinear Equation

The procedure described in the previous section can be extended to nonlinear scalar equations. This is illustrated here using a version of the inviscid Burgers'
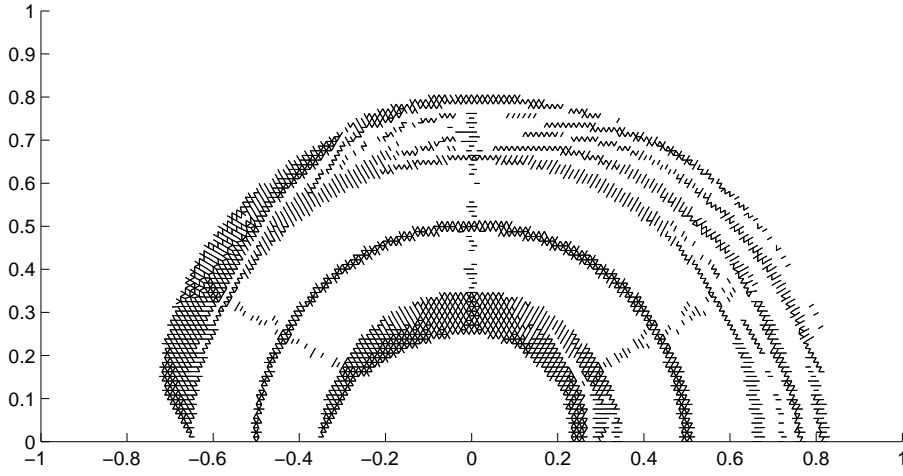
Fig. 16. Illustration of the mesh edges along which the solution value at the quadrature point has been modified when the Gradient PSI scheme is used, applied to Test Case B.

equation which takes the form

$$u_t + \left(\frac{u^2}{2}\right)_x + u_y = 0 \qquad \text{or} \qquad u_t + \vec{\nabla} \cdot \vec{f} = 0 \qquad (62)$$

where $\vec{f} = (\frac{u^2}{2}, u)^{\mathrm{T}}$. This is slightly more complicated than linear advection because Simpson's rule is no longer accurate enough to evaluate the fluctuation

$$\phi = \oint_{\partial \triangle} \vec{f} \cdot \vec{n} \, d\Gamma \qquad (63)$$

exactly. However, it has already been mentioned that the approach can be applied to any order of reconstruction so, once an appropriate N_q-point quadrature has been chosen, the fluctuation can be written

$$\oint_{\partial \triangle} \vec{f} \cdot \vec{n} \, d\Gamma = \sum_{edges} \left(\sum_{l=1}^{N_q} w^l \, \vec{f}(u^l) \cdot \vec{n} \right)_e \qquad (64)$$

$$= \sum_{edges} \left(\sum_{l=1}^{N_q} w^l \, \vec{f}(\bar{u}^l) \cdot \vec{n} \right)_e$$

$$+ \sum_{edges} \left(\sum_{l=1}^{N_q} w^l \left(\vec{f}(u^l) - \vec{f}(\bar{u}^l) \right) \cdot \vec{n} \right)_e ,$$

where w^l are the quadrature weights and $\bar{u}$ is the value of the linear interpolant of u at that point along the given edge.

The polynomial interpolant is limited as before, but now the advection velocity also depends on this value. However, the only new aspect of the above equation

34

is that it contains a difference in the flux rather than the dependent variable. This flux difference can be treated exactly as it would in a standard finite volume scheme (when an appropriate Roe linearisation exists), writing it in the form

$$\vec{f}(u) - \vec{f}(\bar{u}) = \widetilde{\frac{\partial \vec{f}}{\partial u}}(u - \bar{u}) = \tilde{\vec{\lambda}}(u - \bar{u}),$$ (65)

where $\tilde{\vec{\lambda}} = ((u + \bar{u})/2, 1)^T$. One consequence of this is that the time-step limit now depends on the solution, but as long as this solution is bounded, positivity can be maintained. Thus, for this nonlinear scalar equation, the fluctuation can be written in the form

$$\oint_{\partial\triangle} \vec{f} \cdot \vec{n} \, d\Gamma = \phi^{LO} + \sum_{edges} \left(\sum_{l=1}^{N_q} w^l (u^l - \bar{u}^l) \tilde{\vec{\lambda}} \cdot \vec{n} \right)_e,$$ (66)

which has exactly the same structure as the fluctuation given in (34). The interpolant is limited in precisely the same manner and, because $\vec{\lambda} = \frac{\partial \vec{f}}{\partial u}$ is bounded, this again leads to a positive scheme when an appropriate limit is imposed on the time-step, cf. (51) and (55). In fact, the procedure is the same as for linear advection, except that additional quadrature points are used in the integration along each (sub)cell edge. The limiting is carried out as in (41) and (42) and the minimum value of α taken along the edge, to remain consistent with the limited correction taking the form $\delta u'_{i_1 i_2}(\vec{x}) = C'_{i_1 i_2} \delta u_{i_1 i_2}(\vec{x})$.
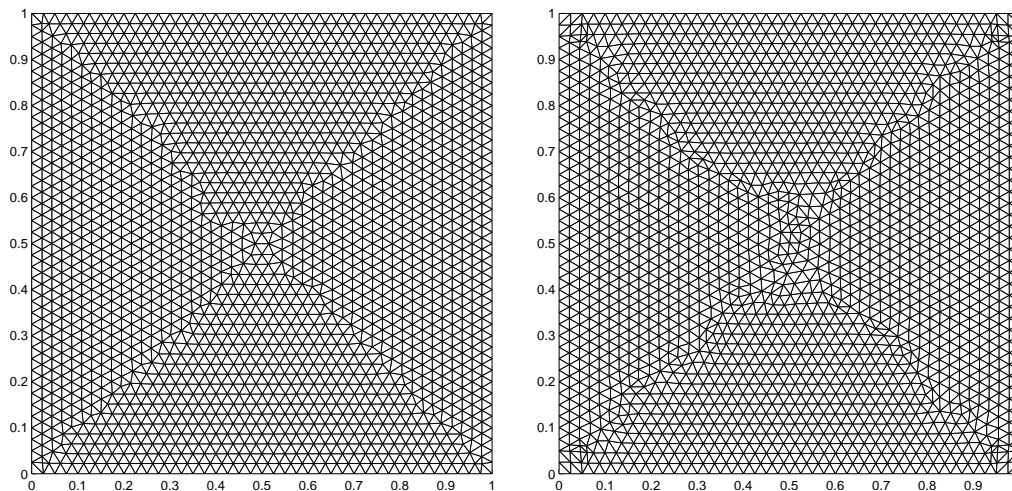

Fig. 17. The meshes used for the inviscid Burgers' equation results obtained with the N, PSI and gradient recovery schemes (left) and the submesh reconstruction schemes (right).

6.1 Results

The two-dimensional inviscid Burgers' equation (62) is approximated here on uniform unstructured triangular meshes covering the domain $[0,1] \times [0,1]$, as shown in Figure 17, with boundary conditions on the inflow boundaries (left, right and bottom) given by

$$u(x,y) = 1.5 - 2x. \tag{67}$$

This problem has the exact solution

$$u(x,y) = \begin{cases} -0.5 & \text{if } y \geq 0.5 \text{ and } -2(x-0.75) + (y-0.5) \leq 0 \\ 1.5 & \text{if } y \geq 0.5 \text{ and } -2(x-0.75) + (y-0.5) \geq 0 \\ \max\left(-0.5, \min\left(1.5, \frac{x-0.75}{y-0.5}\right)\right) & \text{otherwise.} \end{cases} \tag{68}$$

The solutions to this problem obtained using the 6 approaches described in this paper are shown in Figure 18, which shows contour plots of each of the solutions, and Figure 19, which compares magnified sections of one-dimensional slices through the profile along the lines $x = 0.25, 0.5, 0.75, 1.0$. The new limiting procedure again removes completely the unphysical oscillations, but the improvement in accuracy obtained by using the higher order representation is less clear than it was for the linear advection equation.

The differences are most clearly illustrated in the magnified one-dimensional profiles shown in Figure 19. In all cases the N scheme is the most diffusive and the two non-positive schemes give oscillations close to the discontinuity. The two positive high order schemes are also both better than the PSI scheme, but the differences are fairly small, particularly close to the discontinuity, and there is little to choose between the two schemes using a limited quadratic representation. The convergence histories shown in Figure 20 show that all of the methods except the Abgrall-Roe scheme again converge to machine accuracy. The only difference now is that all of the positive schemes converge at a very similar rate, nearly three times as fast as the LDA-based scheme.

These schemes have been applied to a range of initial profiles for the inviscid Burgers' equation (62) and the results suggest that for smooth solutions, the original Caraeni scheme is the best method to use, but when discontinuities appear one of the positive schemes should be applied, the higher order methods giving slightly sharper shocks than the PSI scheme.
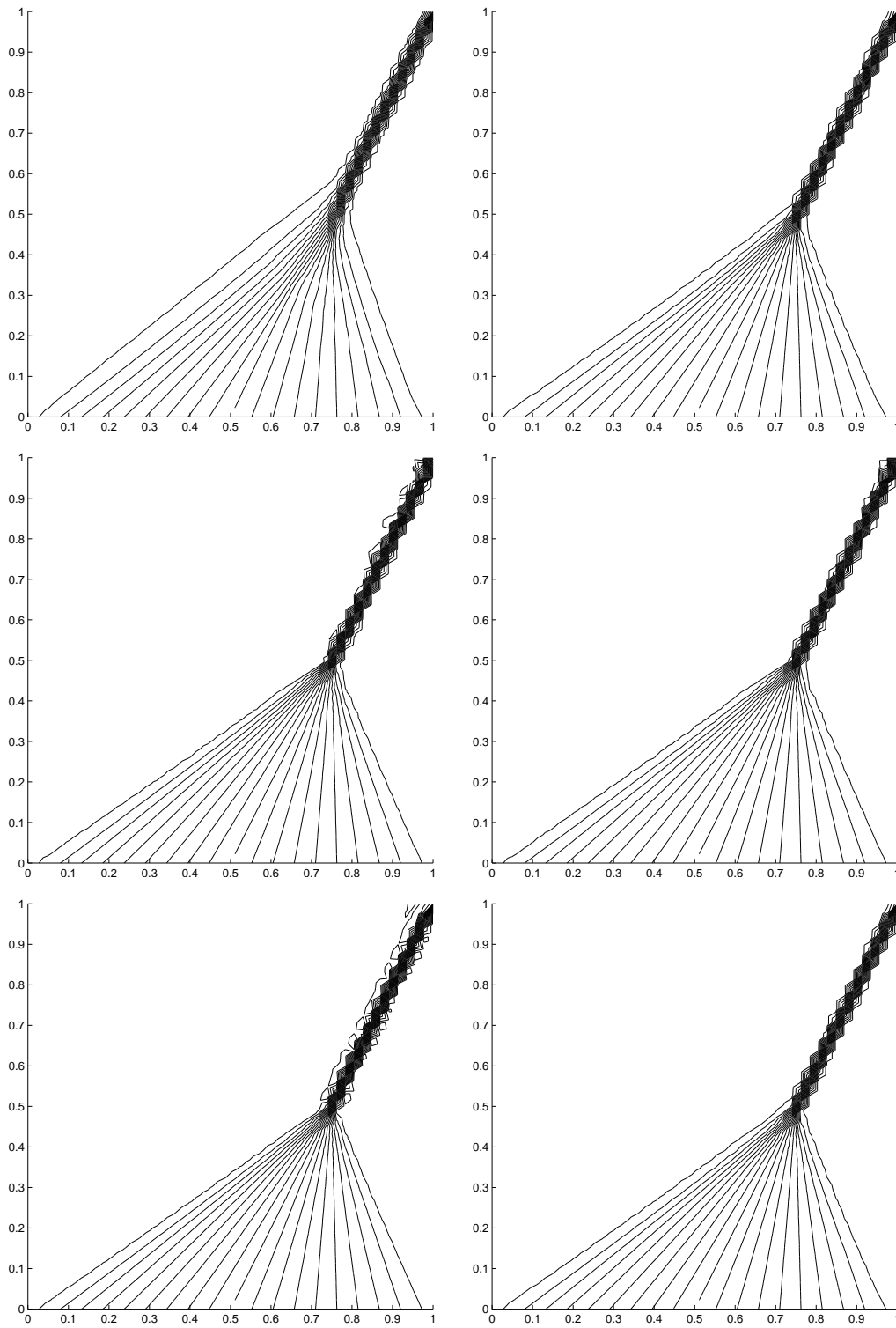
Fig. 18. Solutions to the inviscid Burgers' equation test case for the N scheme (top left), the PSI scheme (top right), the Abgrall-Roe scheme (middle left), the Submesh PSI scheme (middle right), the Caraeni scheme (bottom left) and Gradient PSI scheme (bottom right).
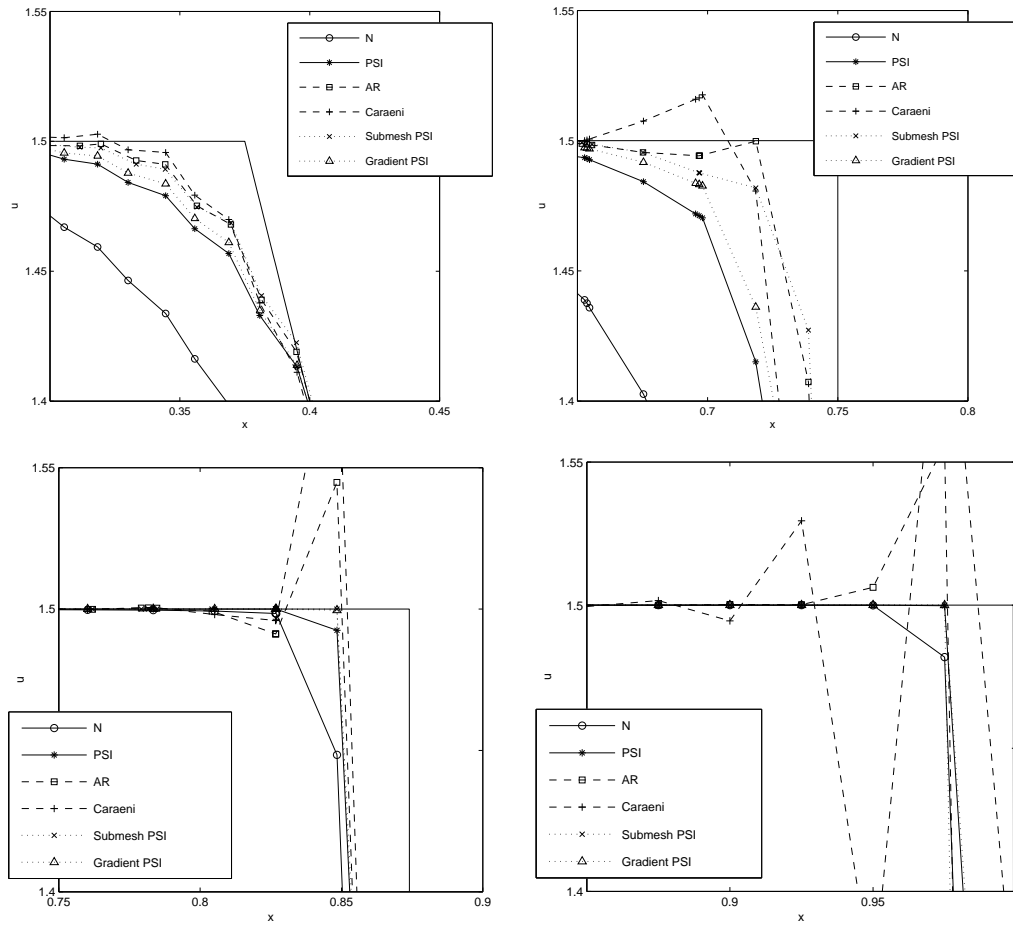
37

Fig. 19. Slices through the solutions to the inviscid Burgers' equation test case for the 6 schemes along $x = 0.25$ (top left), $x = 0.5$ (top right), $x = 0.75$ (bottom left) and $x = 1.0$ (bottom right), magnifying the region where the solution drops from $u = 1.5$. The solid lines with no symbols give the exact solution.
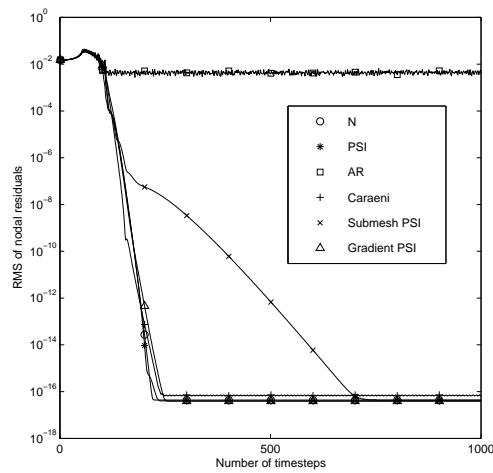

Fig. 20. Convergence histories for each of the methods on the meshes shown in Figure 17 for the inviscid Burgers' equation test case.

7 Conclusion

The issues associated with the construction of a conservative fluctuation splitting scheme which is both higher than second order accurate and positive have been discussed, leading to the proposal of a modification which can be applied to any of the existing high order schemes to impose positivity on them. It is then shown that this produces a family of schemes which can be used to accurately approximate solutions of linear and nonlinear scalar conservation laws without creating any unphysical oscillations in the solution. The effectiveness of the approach is illustrated for a third order scheme on unstructured triangular meshes in two space dimensions. The theory extends straightforwardly to three space dimensions and to arbitrarily high order schemes but, as with the second order PSI scheme on which the new approaches are based, it is not yet clear how best to apply this approach to nonlinear systems of equations. The extension to time-dependent problems is also a challenging issue which is currently being investigated.

The nature of the new scheme, which first produces a limited interpolant of the dependent variable edge-by-edge, also provides the basis for a p-refinement strategy for fluctuation splitting schemes. This could be implemented by using a modification of the proposed approach to smooth the discontinuities created by having a difference in the order of the representation between cells on either side of a given edge, giving a continuous representation and leading to a limited fluctuation and a conservative scheme (which could also be positive if desired). Alternatively, discontinuities could be allowed and the limiting procedure applied to the interpolant could be modified to detect discontinuities in the underlying flow and ensure that any such discontinuities in the representation are kept away from these regions. The resulting scheme would not be conservative, but the allowed discrepancies should not cause significant errors in the calculations.

8 Acknowledgements

References

[1] Abgrall R., "Toward the ultimate conservative scheme: Following the quest", *J. Comput. Phys.*, vol. 167, 2001, p. 277–315.

[2] Abgrall R., "Construction of high order residual distribution schemes for scalar problems. Preliminary results for systems.", in *High Order Discretization Methods*, von Karman Institute for Fluid Dynamics, 2005.

[3] Abgrall R., "Essentially non-oscillatory residual distribution schemes for hyperbolic problems", *J. Comput. Phys.*, vol. 214, 2006, p. 773–808.

[4] Abgrall R., Andrianov N., Mezine M., "Towards very high order accurate schemes for unsteady convection problems on unstructured meshes", *Int. J. Numer. Meth. Fluids*, vol. 47, 2005, p. 679–691.

[5] Abgrall R., Barth T.J., "Residual distribution schemes for conservation laws via adaptive quadrature", *SIAM J. Sci. Comput.*, vol. 24, 2002, p. 732–769.

[6] Abgrall R., Mezine M., "Construction of second order accurate monotone and stable residual distributive schemes: the unsteady case", *J. Comput. Phys.*, vol. 188, 2003, p. 16–55.

[7] Abgrall R., Mezine M., "Residual distribution scheme for steady problems", in *Novel Methods for Solving Convection Dominated Systems*, VKI LS 2003–05, von Karman Institute for Fluid Dynamics, 2003.

[8] Abgrall R., Roe P.L., "High order fluctuation schemes on triangular meshes", *J. Sci. Comput.*, vol. 19, 2003, p. 3–36.

[9] Caraeni D., "A third order residual distribution method for steady/unsteady simulations: formulation and benchmarking using LES", in *High Order Discretization Methods*, von Karman Institute for Fluid Dynamics, 2005.

[10] Caraeni D., Caraeni M., Fuchs L., "A parallel multidimensional upwind algorithm for LES", AIAA paper 2001–2547, 15th AIAA CFD Conference, Anaheim, 2001.

[11] Caraeni D., Fuchs L., "Compact third-order multidimensional upwind scheme for Navier-Stokes simulations", *Theor. Comp. Fluid Dyn.*, vol. 15, 2002, p. 373–401.

[12] Deconinck H., Sermeus K., Abgrall R., "Status of multidimensional upwind residual distribution schemes and applications in aeronautics", AIAA paper 2000–2328, 2000.

[13] Deconinck H., Struijs R., Bourgois G., Roe P.L., "High resolution shock capturing cell vertex advection schemes for unstructured grids", in *Computational Fluid Dynamics*, VKI LS 1994–05, von Karman Institute for Fluid Dynamics, 1994.

[14] FERRANTE A., "Solution of the unsteady Euler equations using residual distribution and flux corrected transport", VKI PR 1997–08, von Karman Institute for Fluid Dynamics, 1997.

[15] HUBBARD M.E., "Non-oscillatory third order fluctuation splitting schemes", in Finite Volumes for Complex Applications IV, 2005, p. 377–387.

[16] HUBBARD M.E., ROE P.L., "Compact high resolution algorithms for time dependent advection on unstructured grids", *Int. J. Numer. Meth. Fluids*, vol. 33, 2000, p. 711–736.

[17] HUBBARD M.E., LAIRD A.L., "Achieving high-order fluctuation splitting schemes by extending the stencil", *Comput. Fluids*, vol. 34, 2005, p. 443–459.

[18] LÖHNER R., MORGAN K., VAHDATI K., BORIS J.P., BOOK D.L., "FEM-FCT: combining unstructured grids with high resolution", *Communications in Applied Numerical Methods*, vol. 4, 1988, p. 717–729.

[19] PAILLÈRE H., DECONINCK H., VAN DER WEIDE E., "Upwind residual distribution methods for compressible flow: an alternative to finite volume and finite element methods", in *C*omputational Fluid Dynamics, VKI LS 1997–02, von Karman Institute for Fluid Dynamics, 1997.

[20] QUINTINO T., RICCHIUTO M., CSÍK Á., DECONINCK H., POEDTS S., "Conservative multidimensional upwind residual distribution schemes for arbitrary finite elements", in Computational Fluid Dynamics, 2002, p. 88–93.

[21] RICCHIUTO M., CSÍK Á., DECONINCK H., "Residual distribution for general time-dependent conservation laws", *J. Comput. Phys.*, vol. 209, 2005, p. 249–289.

[22] SIDILKOVER D., ROE P.L., "Unification of some advection schemes in two dimensions", ICASE Report 95-10, 1995.

[23] STRUIJS R., "A multi-dimensional upwind discretization method for the Euler equations on unstructured grids", Ph.D. Thesis, The University of Delft, The Netherlands, 1994.

[24] ZALESAK S.T., "Fully multidimensional flux-corrected transport algorithms for fluids", *J. Comput. Phys.*, vol. 31, 1979, p. 335–362.