

Positivity-Preserving Flux Difference Splitting Schemes

Bernard Parent*

A positivity-preserving variant of the Roe flux difference splitting method is here proposed. Positivity-preservation is attained by modifying the Roe scheme such that the coefficients of the discretization equation become positive, with a coefficient considered positive if all its eigenvalues are positive and if its eigenvectors correspond to those of the flux Jacobian. Because the modification does not alter the wave speeds at the interface, the appealing attributes of the Roe flux difference splitting schemes are retained, such as high-resolution capture of discontinuous waves, low amount of artificial dissipation within viscous layers, and ease of convergence to steady-state. The proposed flux function is advantaged over previous positivity-preserving variants of the Roe method by being written in general matrix form and hence by being readily deployable to arbitrary systems of conservation laws. The stencils are extended to second-order accuracy through a newly-derived positivity-preserving total-variation-diminishing limiting process that is applied to the characteristic variables and that yields positive coefficients. Also derived is a positivity-preserving restriction on the time step for flux difference splitting schemes that is shown to depart significantly from the CFL condition in regions with high property gradients.

1. Introduction

ORIGINALLY published more than three decades ago, the Roe flux difference splitting scheme [1, 2] remains today one of the most used methods to discretize the convection derivatives within fluid flow systems of conservation laws. The lasting popularity of the Roe scheme lies in it having the following three properties: (i) it is monotonicity-preserving, (ii) it introduces minimal dissipation within viscous layers and discontinuities, and (iii), it is written in general matrix form. Indeed, when arithmetic averaging instead of Roe averaging is used to determine the Jacobian at the interface, the Roe flux is written in general matrix form because it is function only of the flux vector, of the vector of conserved variables, and of the eigenvalues and eigenvectors of the flux Jacobian. This makes it possible to deploy the Roe scheme, without modification, to arbitrary systems of conservation laws. Other commonly-used flux discretization approaches may have one or two of the properties just listed, but not all three. For instance, the Godunov exact Riemann solver [3], the HLLC approximate Riemann solver [4], and the AUSM method [5] are not written in general matrix form, while the Steger-Warming flux vector splitting method [6] and the HLL approximate Riemann solver [7] suffer from excessive dissipation within viscous layers.

The Roe scheme has nonetheless one major disadvantage over competing methods: it is not positivity-preserving. Positivity-preservation refers to the capability of a discretization stencil to maintain the positivity of the determinative properties, with the latter being the properties that must be positive for the solution to be physically-permissible. For instance, the determinative properties associated with the Euler equations are the density and the temperature; the determinative properties associated with the multispecies Favre-averaged Navier-Stokes equations would further include the partial densities, the turbulence kinetic energy, and its dissipation rate. The Roe scheme is well-known not to maintain the positivity of the determinative properties, with the effect that negative densities, temperatures or turbulence kinetic energies appear occasionally in the solution. This can be remedied through a “clipping” of the determinative properties after each iteration to ensure that they remain physically meaningful. But such is an undesirable fix because, when solving time-accurate cases, this leads to a loss of conservation which can induce substantial error within the solution, and when solving steady-state cases, this further leads to convergence issues often preventing a converged solution altogether.

Despite the considerable interest over the years to develop a positivity-preserving variant of the Roe flux function, limited success has been reported to date. For instance, in Ref. [8], Einfeldt et al show that a modification of the averaging process at

*Faculty Member, Dept. of Aerospace Engineering, Pusan National University, Busan 609-735, Korea, <http://www.bernardparent.com>.

the interfaces cannot lead to positivity-preservation. That is, it is not possible to make the Roe scheme positivity-preserving by substituting the Roe average by an arithmetic average or any other type of averaging process. More success is reported in Refs. [9, 10], in which Dubroca proposes a new family of Roe matrices: the eigenvectors and eigenvalues are not obtained from the convective flux Jacobian as is usually done, but rather from a matrix that is as close as possible to the flux Jacobian while resulting in a positivity-preserving discretization stencil. However, the approach proposed by Dubroca is limited to the one-dimensional Euler equations and it is not clear how it can be extended to the Euler equations in multiple dimensions, let alone to other systems of conservation laws including the transport equations of nitrogen vibrational energy, turbulence kinetic energy, partial densities, etc.

Meanwhile, some progress in positivity-preserving theory for fluid flow systems of conservation laws has been reported that is of particular interest. In Ref. [11], a new approach named the "rule of the positive coefficients" is proposed to craft positivity-preserving discretization stencils. Using the rule of the positive coefficients, Parent derives in Ref. [12] a new set of total-variation-diminishing (TVD) limiters which are advantaged over the conventional limiters by being positivity-preserving. The positivity-preserving TVD stencils proposed by Parent are particularly appealing for two reasons: (i) they do not introduce more dissipation than the conventional TVD stencils, and (ii) they are written in general matrix form. That is, the stencils depend solely on the vector of conserved variables and on the flux Jacobian and can hence be readily deployed to any system of conservation laws. However, because the method is a second-order extension of the Steger-Warming scheme, it suffers from the same limitations as the latter and hence introduces excessive dissipation within viscous layers compared to Roe flux functions.

The goal of this paper is to use the rule of the positive coefficients to craft a novel variant of the Roe flux function that is positivity-preserving and that is written in general matrix form (i.e. only function of the flux vector, of the vector of conserved variables, and of the flux Jacobian). The proposed method is hence advantaged over previous positivity-preserving variants of the Roe method by being readily extendable to any system of conservation laws, and is advantaged over previous positivity-preserving Steger-Warming schemes by being capable to capture viscous layers with high resolution.

This paper is organized as follows. First, an outline is given of the rule of the positive coefficients including its possible limitations. Then, the Roe scheme is recast in a form similar to the Steger-Warming flux vector splitting method in order to ease the task of attaining positivity-preserving stencils. This is followed by the outline of novel positivity-preserving variants of the Roe flux functions written in general matrix form, and of a novel positivity-preserving condition on the time step applicable to flux difference splitting schemes. Lastly, several test cases solving the Euler equations and the Navier-Stokes equations are presented to assess the positivity-preserving and resolution capabilities of the proposed stencils compared to the conventional Roe methods.

2. Rule of the Positive Coefficients

The rule of the positive coefficients can be summarized as follows. Consider a discretization equation updating the vector of conserved variables U at the i th node as follows:

$$C_i^{n+1}U_i^{n+1} = C_iU_i + C_{i+1}U_{i+1} + C_{i-1}U_{i-1} + \dots \quad (1)$$

where i is the grid index, n the iteration count, and where C_i , C_{i+1} , etc are the discretization coefficients (in matrix form). In Ref. [11], it is demonstrated that a discretization stencil maintains the positivity of the determinative properties as long as all the coefficients within the discretization equation are positive. A "determinative property" here refers to a property that must necessarily be positive to be physically meaningful (such as the pressure, the density, etc), and a "positive coefficient" refers to a coefficient with all-positive eigenvalues and with the same eigenvectors as those of the respective flux Jacobian:

$$C_i = L_i^{-1}D_i^+L_i \quad , \quad C_{i-1} = L_{i-1}^{-1}D_{i-1}^+L_{i-1} \quad , \quad C_{i+1} = L_{i+1}^{-1}D_{i+1}^+L_{i+1} \quad , \quad \dots \quad (2)$$

where L and L^{-1} correspond to the left and right eigenvectors of the flux Jacobian and D^+ to a diagonal matrix with all-positive diagonal elements. Although the proof presented in Ref. [11] is limited to the perfect-gas Euler equations, there are indications that the rule of the positive coefficients remains valid for other physical models. For instance, when solving a mixture of gases, the eigenvectors and eigenvalues associated with the additional species conservation equations have a similar form as those associated with the total mass conservation equation. Then, following the same steps as in the Appendix of Ref. [11], it can be shown that the positivity of each species density is conserved as long as the coefficients within the discretization equation are all positive. Similar arguments can also be made for the additional transport equations related to turbulence modeling: when a two-equation turbulence model is solved in conjunction with the mass, momentum, and energy transport equations, the turbulence kinetic energy and its dissipation rate can be shown to remain positive should the rule of the positive coefficients

be enforced. What is more challenging to demonstrate, however, is whether the internal energy (and hence, the pressure and temperature) would remain positive when the system of conservation laws include these additional transport equations and/or when the gas is non perfect. The most that can be stated at this stage is that no negative pressures and temperatures were encountered in doing numerous numerical experiments of a real multispecies gas including a two-equation turbulence model as long as the discretization equation obeyed the rule of the positive coefficients. While this does not constitute a proof, it does shed hope that the rule of the positive coefficients is not limited to the perfect-gas Euler equations but also applies to other sets of equations commonly encountered in fluid mechanics, and perhaps even in plasmadynamics.

3. Recast of Flux Difference Splitting Schemes in Flux Vector Splitting Form

The first-order Roe flux difference splitting schemes is here rewritten (without loss of generality) in a form similar to the Steger-Warming flux vector splitting methods. The reason for this will become apparent in subsequent sections when some positivity-preserving variants of the Roe flux functions are derived. To do so, first note that the Roe flux difference splitting scheme yields a flux at the interface of the form [1]:

$$F_{i+1/2} = \frac{1}{2}F_i + \frac{1}{2}F_{i+1} - \frac{1}{2}|A|_{i+1/2}(U_{i+1} - U_i) \quad (3)$$

where F is the convective flux, U the vector of conserved variables, and $|A|$ the Roe matrix equivalent to:

$$|A| \equiv L^{-1}|\Lambda|L \quad (4)$$

where L is the left eigenvector matrix, L^{-1} is the right eigenvector matrix, Λ is the eigenvalue matrix of the convective flux Jacobian $A \equiv \partial F/\partial U$, and $|\Lambda|$ is the absolute value of the eigenvalue matrix obtained from Λ by taking the absolute value of all elements. To ensure that the scheme does not induce non-physical phenomena, it is necessary to correct the eigenvalues part of the matrix $|A|$ as follows:

$$|\Lambda|_{r,r} \rightarrow \sqrt{|\Lambda|_{r,r}^2 + \delta a^2} \quad (5)$$

with a the speed of sound and δ a user-defined parameter. Then recast Eq. (3) in Steger-Warming flux vector splitting form as follows:

$$F_{i+1/2} = L_i^{-1}\Lambda_i^+L_iU_i + L_{i+1}^{-1}\Lambda_{i+1}^-L_{i+1}U_{i+1} \quad (6)$$

where Λ_i^\pm are diagonal matrices that will be determined below. It is noted that because the eigenvectors are obtained from the Jacobian A which is defined as $A \equiv \partial F/\partial U$, it follows that Eq. (6) necessarily implies that the set of equations must be homogenous of degree 1. But this is not a particular source of concern: not only do the perfect-gas Euler equations have this property, but so do various other sets of equations used to solve compressible fluid flow (such as the multi-species real-gas Euler equations, the Favre-averaged Navier-Stokes equations, etc). Further, should the set of equations not be homogeneous of degree 1, the methods developed herein can still be used provided the matrix A is redefined such that $AU \equiv F$.

By imposing the condition that the flux in Steger-Warming form outlined in Eq. (6) must be equal to the flux in Roe form outlined in Eq. (3), it is apparent that the following two equalities must hold:

$$L_i^{-1}\Lambda_i^+L_iU_i = \frac{1}{2}F_i + \frac{1}{2}|A|_{i+1/2}U_i \quad (7)$$

$$L_{i+1}^{-1}\Lambda_{i+1}^-L_{i+1}U_{i+1} = \frac{1}{2}F_{i+1} - \frac{1}{2}|A|_{i+1/2}U_{i+1} \quad (8)$$

The latter essentially consist of the definitions of the diagonal matrices Λ_i^\pm . Then, note that for a system of conservation laws that is homogeneous of degree one, the convective flux can be written as $F_i = A_iU_i = L_i^{-1}\Lambda_iL_iU_i$. After making the latter substitution and multiplying all terms in Eq. (7) by L_i , the following is obtained:

$$\Lambda_i^+L_iU_i = \frac{1}{2}\Lambda_iL_iU_i + \frac{1}{2}L_i|A|_{i+1/2}U_i \quad (9)$$

Rewrite the latter in tensor form and divide all terms by $[L_iU_i]_r$:

$$[\Lambda_i^+]_{r,r} = \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r} \quad (10)$$

Starting from Eq. (8) and following a similar procedure as outlined above, an expression for Λ^- can be found:

$$[\Lambda_{i+1}^-]_{r,r} = \frac{1}{2} [\Lambda_{i+1}]_{r,r} - \frac{1}{2} \frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r} \quad (11)$$

In the latter 2 equations, the last term on the RHS can be seen to involve a division by the characteristic variables $[L_i U_i]_r$. It may be argued that this could lead to a division by zero in solving the Euler equations in multiple dimensions. This is not a cause for concern, however, because there exists a set of eigenvectors and eigenvalues for the 2D and 3D Euler equations that is such that the characteristic variables never become zero (see the Appendix of Ref. [12] for instance).

It is emphasized that, thus far, the Roe scheme has not been modified, but has simply been recast in a different form, similar to the one used to express the Steger-Warming flux vector splitting method. Indeed, by substituting Eqs. (10) and (11) into Eq. (6), one would obtain exactly the Roe scheme.

4. Proposed Positivity-Preserving Flux Difference Splitting Schemes

Some novel first-order flux functions are now proposed that retain the appealing features of the Roe scheme while being positivity-preserving. The schemes are derived for a 1D hyperbolic system of conservation laws, which can be written in discrete form as:

$$\frac{U_i^{n+1} - U_i}{\Delta t} + \frac{F_{i+1/2} - F_{i-1/2}}{\Delta x} = 0 \quad (12)$$

where U is the vector of conserved variables, F is the convective flux, n denotes the time level, and i is the grid index along x . Although the schemes are here derived in 1D, they can be extended to 2D and 3D through dimensional splitting while remaining positivity-preserving.

The first-order Roe scheme can be made positivity-preserving by modifying the eigenvalues Λ^\pm outlined above in Eqs. (10)-(11) such that the discretization equation conforms to the rule of the positive coefficients. To determine the conditions on Λ^\pm that result in the discretization equation satisfying the rule of the positive coefficients, first substitute the interface flux outlined in Eq. (6) into the discrete equation (12). This yields:

$$\frac{U_i^{n+1} - U_i}{\Delta t} + \frac{L_i^{-1} \Lambda_i^+ L_i U_i + L_{i+1}^{-1} \Lambda_{i+1}^- L_{i+1} U_{i+1}}{\Delta x} - \frac{L_{i-1}^{-1} \Lambda_{i-1}^+ L_{i-1} U_{i-1} + L_i^{-1} \Lambda_i^- L_i U_i}{\Delta x} = 0 \quad (13)$$

The latter can then be rewritten in standard discretization equation form as follows:

$$C_i^{n+1} U_i^{n+1} = C_{i-1} U_{i-1} + C_i U_i + C_{i+1} U_{i+1} \quad (14)$$

with the discretization coefficients defined as:

$$C_i^{n+1} \equiv \frac{\Delta x}{\Delta t} L_i^{-1} I L_i \quad (15)$$

$$C_i \equiv L_i^{-1} \left(\frac{\Delta x}{\Delta t} I - \Lambda_i^+ + \Lambda_i^- \right) L_i \quad (16)$$

$$C_{i-1} \equiv L_{i-1}^{-1} \Lambda_{i-1}^+ L_{i-1} \quad (17)$$

$$C_{i+1} \equiv -L_{i+1}^{-1} \Lambda_{i+1}^- L_{i+1} \quad (18)$$

According to the rule of the positive coefficients, a discretization stencil is positivity-preserving if all the coefficients are positive [11]. For a coefficient to be considered positive, two conditions must be fulfilled. First, its eigenvectors must correspond to those of the convective flux Jacobian evaluated at the corresponding node. This can be verified to be the case for all the above coefficients. Second, its eigenvalues must all be positive. This is always true for the coefficient C_i^{n+1} , and can become true for the coefficient C_i should the time step be sufficiently small. The other two coefficients, C_{i-1} and C_{i+1} , are guaranteed to be positive only if all the diagonal elements within the matrix Λ^+ are positive and if all the diagonal elements within the matrix Λ^- are negative. Unfortunately, this is not the case for the Roe scheme. Indeed, the Λ^+ and Λ^- matrices associated with the Roe flux function (as derived in the previous section) may not yield positive coefficients. This becomes clear when Eqs. (10) and (11) are rewritten as:

$$[\Lambda_{i+1}^-]_{r,r}^{\text{Roe}} = \underbrace{\min \left(0, \frac{1}{2} [\Lambda_{i+1}]_{r,r} - \frac{1}{2} \frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r} \right)}_{\text{positivity-preserving}} + \underbrace{\max \left(0, \frac{1}{2} [\Lambda_{i+1}]_{r,r} - \frac{1}{2} \frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r} \right)}_{\text{not necessarily positivity-preserving}} \quad (19)$$

$$[\Lambda_i^+]_{r,r}^{\text{Roe}} = \underbrace{\max\left(0, \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r}\right)}_{\text{positivity-preserving}} + \underbrace{\min\left(0, \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r}\right)}_{\text{not necessarily positivity-preserving}} \quad (20)$$

Clearly, the last term on the RHS of the latter two equations may result in Λ^\pm matrices that may not conform to the rule of the positive coefficients. To make the Roe scheme positivity-preserving the Λ^+ eigenvalues presented in Eq. (20) must be modified to be always positive, and the Λ^- eigenvalues presented in Eq. (19) must be modified to be always negative. Then, the coefficients of the discretization equation would respect the rule of the positive coefficients, and the scheme would become positivity-preserving. One way this can be accomplished is by moving the negative terms from Λ^+ to Λ^- and by moving the positive terms from Λ^- to Λ^+ :

$$[\Lambda_{i+1}^-]_{r,r} = \underbrace{\min\left(0, \frac{1}{2}[\Lambda_{i+1}]_{r,r} - \frac{1}{2}\frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r}\right)}_{\text{positivity-preserving}} + \underbrace{\min\left(0, \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r}\right)}_{\text{positivity-preserving}} \quad (21)$$

$$[\Lambda_i^+]_{r,r} = \underbrace{\max\left(0, \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r}\right)}_{\text{positivity-preserving}} + \underbrace{\max\left(0, \frac{1}{2}[\Lambda_{i+1}]_{r,r} - \frac{1}{2}\frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r}\right)}_{\text{positivity-preserving}} \quad (22)$$

The approach proposed here has the advantage of not altering the Roe wave speed at the interface: the amount of wave speed lost on the left node corresponds to the amount of wave speed gained on the right node and vice versa.

A positivity-preserving variant of the Roe scheme can hence be obtained by substituting the latter Λ^+ and Λ^- eigenvalues in the flux at the interface outlined in Eq. (6):

$$F_{i+1/2} = L_i^{-1}G_i^+ + L_{i+1}^{-1}G_{i+1}^- \quad (23)$$

with $G^\pm \equiv \Lambda^\pm LU$. As will be shown in Section 8 below through some test cases, the latter flux function yields results that are very close to those obtained with the conventional Roe method within shocks, expansion fans, and contact discontinuities, but introduces slightly more dissipation within high-Reynolds-number laminar viscous layers.

4.1. Iterative Form

To reduce the dissipation within viscous layers, an alternative approach is here proposed to determine the positive and negative eigenvalues that result in a closer agreement with the original Roe method, albeit at the expense of algorithm complexity. This is accomplished by first substituting the positive and negative eigenvalues outlined in Eqs. (10) and (11) into the Roe flux at the interface Eq. (6). After some reformatting, the following is obtained:

$$F_{i+1/2}^{\text{Roe}} = \underbrace{L_i^{-1}Y_i^+L_iU_i + L_{i+1}^{-1}Z_{i+1}^-L_{i+1}U_{i+1}}_{\text{positivity-preserving}} + \underbrace{L_i^{-1}Y_i^-L_iU_i + L_{i+1}^{-1}Z_{i+1}^+L_{i+1}U_{i+1}}_{\text{not necessarily positivity-preserving}} \quad (24)$$

with the diagonal matrices Y^\pm and Z^\pm set equal to:

$$[Y_i^-]_{r,r} = \min\left(0, \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r}\right) \quad (25)$$

$$[Z_{i+1}^-]_{r,r} = \min\left(0, \frac{1}{2}[\Lambda_{i+1}]_{r,r} - \frac{1}{2}\frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r}\right) \quad (26)$$

$$[Y_i^+]_{r,r} = \max\left(0, \frac{1}{2}[\Lambda_i]_{r,r} + \frac{1}{2}\frac{[L_i|A|_{i+1/2}U_i]_r}{[L_iU_i]_r}\right) \quad (27)$$

$$[Z_{i+1}^+]_{r,r} = \max\left(0, \frac{1}{2}[\Lambda_{i+1}]_{r,r} - \frac{1}{2}\frac{[L_{i+1}|A|_{i+1/2}U_{i+1}]_r}{[L_{i+1}U_{i+1}]_r}\right) \quad (28)$$

Thus far, the Roe scheme has not been modified. But the terms that are positivity-preserving have been separated from those that are not necessarily positivity-preserving. Indeed, the first two terms on the RHS of Eq. (24) are positivity-preserving because they result in discretization coefficients that are positive (i.e., with positive eigenvalues and the same eigenvectors as those of

the flux Jacobian) or that can become positive for a small enough time step. On the other hand, the last two terms on the RHS of Eq. (24) are not positivity-preserving because they may result in one or more of the discretization coefficients having one or more negative eigenvalues.

It follows that one way that the stencil could be made positivity-preserving is simply by dropping the last two terms on the RHS of Eq. (24). But, by doing so, the Roe scheme would be modified substantially and some of its desirable attributes may be lost in the process. For this reason, instead of discarding the two terms that are not necessarily positivity-preserving, let us recast them into new terms, some of which being guaranteed to be positivity-preserving. This can be accomplished by first defining the diagonal matrices R and S such that the following two statements hold:

$$L_{i+1}^{-1} R_{i+1} L_{i+1} U_{i+1} \equiv L_i^{-1} Y_i^- L_i U_i \quad (29)$$

$$L_i^{-1} S_i L_i U_i \equiv L_{i+1}^{-1} Z_{i+1}^+ L_{i+1} U_{i+1} \quad (30)$$

Then, using the latter two definitions, the flux at the interface, Eq. (24), can be rewritten as:

$$F_{i+1/2}^{\text{Roe}} = +L_i^{-1} (Y_i^+ + S_i) L_i U_i + L_{i+1}^{-1} (Z_{i+1}^- + R_{i+1}) L_{i+1} U_{i+1} \quad (31)$$

where the matrices R and S can be obtained in terms of the other matrices by multiplying both sides of Eqs. (29) and (30) by L_{i+1} , writing in tensor form, and then isolating R and S :

$$[R_{i+1}]_{r,r} = \frac{[L_{i+1} L_i^{-1} Y_i^- L_i U_i]_r}{[L_{i+1} U_{i+1}]_r} \quad (32)$$

$$[S_i]_{r,r} = \frac{[L_i L_{i+1}^{-1} Z_{i+1}^+ L_{i+1} U_{i+1}]_r}{[L_i U_i]_r} \quad (33)$$

Now, let us split again the terms within the Roe flux in terms of positivity-preserving fluxes and not-necessarily-positivity-preserving fluxes. This can be done by rewriting Eq. (31) as:

$$F_{i+1/2}^{\text{Roe}} = \underbrace{L_i^{-1} (Y_i^+)^{m+1} L_i U_i + L_{i+1}^{-1} (Z_{i+1}^-)^{m+1} L_{i+1} U_{i+1}}_{\text{positivity-preserving}} + \underbrace{L_i^{-1} (Y_i^-)^{m+1} L_i U_i + L_{i+1}^{-1} (Z_{i+1}^+)^{m+1} L_{i+1} U_{i+1}}_{\text{not necessarily positivity-preserving}} \quad (34)$$

where the superscript m is an iteration counter such that $(\cdot)^{m+1}$ refers to an update of the properties (\cdot) . Then, for Eq. (34) to be equal to the Roe flux at the interface, Eq. (31), the updated Y^\pm and Z^\pm diagonal matrices must be equal to:

$$[Y_i^-]_{r,r}^{m+1} = \min \left(0, [Y_i^+]_{r,r}^m + [S_i]_{r,r}^m \right) \quad (35)$$

$$[Z_{i+1}^-]_{r,r}^{m+1} = \min \left(0, [Z_{i+1}^-]_{r,r}^m + [R_{i+1}]_{r,r}^m \right) \quad (36)$$

$$[Y_i^+]_{r,r}^{m+1} = \max \left(0, [Y_i^+]_{r,r}^m + [S_i]_{r,r}^m \right) \quad (37)$$

$$[Z_{i+1}^+]_{r,r}^{m+1} = \max \left(0, [Z_{i+1}^-]_{r,r}^m + [R_{i+1}]_{r,r}^m \right) \quad (38)$$

where the notation $(\cdot)^m$ denotes the property (\cdot) at the previous iteration count. Then, after substituting R and S from Eq. (32) and (33) the latter 4 equations become:

$$[Y_i^-]_{r,r}^{m+1} = \min \left(0, [Y_i^+]_{r,r}^m + \frac{[L_i L_{i+1}^{-1} (Z_{i+1}^+)^m L_{i+1} U_{i+1}]_r}{[L_i U_i]_r} \right) \quad (39)$$

$$[Z_{i+1}^-]_{r,r}^{m+1} = \min \left(0, [Z_{i+1}^-]_{r,r}^m + \frac{[L_{i+1} L_i^{-1} (Y_i^-)^m L_i U_i]_r}{[L_{i+1} U_{i+1}]_r} \right) \quad (40)$$

$$[Y_i^+]_{r,r}^{m+1} = \max \left(0, [Y_i^+]_{r,r}^m + \frac{[L_i L_{i+1}^{-1} (Z_{i+1}^+)^m L_{i+1} U_{i+1}]_r}{[L_i U_i]_r} \right) \quad (41)$$

$$[Z_{i+1}^+]_{r,r}^{m+1} = \max \left(0, [Z_{i+1}^-]_{r,r}^m + \frac{[L_{i+1} L_i^{-1} (Y_i^-)^m L_i U_i]_r}{[L_{i+1} U_{i+1}]_r} \right) \quad (42)$$

By performing several iterations $m = 1, 2, 3, \dots$, the latter set of equations essentially transforms (as much as possible) the not-necessarily-positivity-preserving terms into positivity-preserving terms (see Eq. (34)). However, it is noted that when used in conjunction with Eq. (34), the latter expressions for Y^\pm and Z^\pm will yield exactly the Roe scheme independently of how many times the matrices Y^\pm and Z^\pm are updated. To make the scheme positivity-preserving, rewrite the flux at the interface, Eq. (34), as:

$$F_{i+1/2} = L_i^{-1} \Lambda_i^+ L_i U_i + L_{i+1}^{-1} \Lambda_{i+1}^- L_{i+1} U_{i+1} \quad (43)$$

with the positive eigenvalues Λ^+ being set to the sum of the positive wave speeds from both the left and right nodes:

$$\Lambda_i^+ = Y_i^+ + Z_{i+1}^+ \quad (44)$$

and with the negative eigenvalues Λ^- defined as the sum of the negative wave speeds originating from both the left and right nodes:

$$\Lambda_{i+1}^- = Z_{i+1}^- + Y_i^- \quad (45)$$

Compared to the original Roe scheme, the latter flux function does not modify the wave speed at the interface: the wave speed lost by the right node is gained by the left node and vice-versa. Because of this, it retains the desirable attributes of the original method such as low dissipation within viscous layers, high resolution in the vicinity of discontinuities, and ease of convergence to steady-state.

In summary, the "iterative form" of the positivity-preserving Roe scheme presented in this section consists of (i) initializing the Y^\pm and Z^\pm diagonal matrices using Eqs. (25) to (28), (ii) updating the Y^\pm and Z^\pm diagonal matrices through an iterative process by using Eqs. (39) to (42), (iii) determining the positive and negative eigenvalues through Eqs. (44) and (45) using the latest updates of the Y^\pm and Z^\pm matrices, and (iv) determining the flux at the interface as in Eq. (43). It is here recommended to update the Y^\pm and Z^\pm matrices only two times, as further updating the latter matrices seldomly results in a noticeable improvement of the solution while requiring more computing effort.

5. Positivity-Preserving Second-Order FDS Flux Functions

One way the Roe scheme can be extended to second-order accuracy while remaining monotonicity-preserving is through the use of TVD limiters applied to the characteristic variables as proposed by Yee [13, 14, 15]. Commonly denoted as the "Yee-Roe" scheme, such a strategy has enjoyed considerable popularity in solving compressible viscous flows because it is second-order accurate and, like the first-order Roe scheme, it introduces little dissipation in viscous layers, it is monotonicity-preserving, and it converges reliably for a wide variety of flow conditions.

The Yee-Roe flux at the interface can be written as follows:

$$F_{i+1/2} = \underbrace{\frac{1}{2}F_i + \frac{1}{2}F_{i+1} - \frac{1}{2}L_{i+1/2}^{-1}|\Lambda|_{i+1/2}M_{i+1/2}}_{\text{first-order Roe terms}} + \underbrace{\frac{1}{2}L_{i+1/2}^{-1}|\Lambda|_{i+1/2}\Phi_{i+1/2}M_{i+1/2}}_{\text{second-order Yee terms}} \quad (46)$$

where the vector $M_{i+1/2}$ is defined as:

$$M_{i+1/2} \equiv L_{i+1/2}(U_{i+1} - U_i) \quad (47)$$

and where the diagonal elements within the limiter matrix are set as follows:

$$[\Phi_{i+1/2}]_{r,r} = \frac{\text{minmod}([M_{i-1/2}]_r, [M_{i+1/2}]_r, [M_{i+3/2}]_r)}{[M_{i+1/2}]_r} \quad (48)$$

where the minmod function returns the argument with the smallest magnitude if the arguments all share the same sign and zero if the arguments are of mixed signs. A possible division by zero on the RHS of Eq. (48) can be avoided by adding a small constant to the denominator, as is common practice when implementing the Yee-Roe scheme. This does not pose problems when doing computations because the minmod function on the numerator will always be as low in magnitude as the denominator, hence keeping the limiter function bounded.

As was shown above, the first-order Roe terms can be recast in flux vector splitting form with the positive and negative eigenvalue matrices Λ^\pm as outlined in Eqs. (10) and (11). Without loss of generality, Eq. (46) can hence be rewritten in flux-vector splitting form as:

$$F_{i+1/2} = L_i^{-1} \Lambda_i^+ L_i U_i + L_{i+1}^{-1} \Lambda_{i+1}^- L_{i+1} U_{i+1} + \frac{1}{2} L_{i+1/2}^{-1} |\Lambda|_{i+1/2} \Phi_{i+1/2} M_{i+1/2} \quad (49)$$

Now, let us recast the second-order Yee terms in a form that can be easily made positivity-preserving. For this purpose, the second-order terms can be rewritten similarly to the positivity-preserving TVD schemes proposed by Parent for flux vector splitting methods in Ref. [12]. This can be accomplished as follows. First note that the absolute value of the eigenvalue matrix can be rewritten as:

$$|\Lambda|_{i+1/2} = \frac{1}{2} (\Lambda_{i+1/2} + |\Lambda|_{i+1/2}) - \frac{1}{2} (\Lambda_{i+1/2} - |\Lambda|_{i+1/2}) \quad (50)$$

Substitute the latter in the former:

$$F_{i+1/2} = L_i^{-1} \Lambda_i^+ L_i U_i + L_{i+1}^{-1} \Lambda_{i+1}^- L_{i+1} U_{i+1} + \frac{1}{4} L_{i+1/2}^{-1} (\Lambda_{i+1/2} + |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2} - \frac{1}{4} L_{i+1/2}^{-1} (\Lambda_{i+1/2} - |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2} \quad (51)$$

Keep the latter on hold. Now, define the limiter matrices Ψ^\pm such that the following hold:

$$L_i^{-1} \Psi_{i+1/2}^+ (G_{i+1}^+ - G_i^+) \equiv \frac{1}{2} L_{i+1/2}^{-1} (\Lambda_{i+1/2} + |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2} \quad (52)$$

$$L_{i+1}^{-1} \Psi_{i+1/2}^- (G_{i+1}^- - G_i^-) \equiv \frac{1}{2} L_{i+1/2}^{-1} (\Lambda_{i+1/2} - |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2} \quad (53)$$

where G^\pm is defined as:

$$G^\pm \equiv \Lambda^\pm L U \quad (54)$$

After substituting Eq. (52) and Eq. (53) into Eq. (51) the Yee-Roe flux becomes:

$$F_{i+1/2} = L_i^{-1} \Lambda_i^+ L_i U_i + L_{i+1}^{-1} \Lambda_{i+1}^- L_{i+1} U_{i+1} + \frac{1}{2} L_i^{-1} \Psi_{i+1/2}^+ (G_{i+1}^+ - G_i^+) - \frac{1}{2} L_{i+1}^{-1} \Psi_{i+1/2}^- (G_{i+1}^- - G_i^-) \quad (55)$$

where the Ψ^+ diagonal limiter matrix can be obtained by first multiplying all terms within Eq. (52) by L_i , then rewriting in tensor form, and then isolating Ψ^+ :

$$[\Psi_{i+1/2}^+]_{r,r} = \frac{[L_i L_{i+1/2}^{-1} (\Lambda_{i+1/2} + |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2}]_r}{2 [G_{i+1}^+ - G_i^+]_r} \quad (56)$$

Similarly, to obtain an expression for the Ψ^- diagonal limiter matrix, multiply all terms within Eq. (53) by L_{i+1} , rewrite in tensor form, and isolate Ψ^- :

$$[\Psi_{i+1/2}^-]_{r,r} = \frac{[L_{i+1} L_{i+1/2}^{-1} (\Lambda_{i+1/2} - |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2}]_r}{2 [G_{i+1}^- - G_i^-]_r} \quad (57)$$

As is commonly done with other TVD limiter functions, a small user-defined constant is added to the denominator on the RHS of Eqs. (56) and (57) to prevent a division by zero. This is not problematic and does not lead to a significant dependence of the flux on an arbitrary user-defined constant because the limiter function Ψ is eventually multiplied by the same quantity that appears on its denominator when calculating the flux function in Eq. (55).

Thus far, the Yee-Roe flux has not been modified. Indeed, when the eigenvalues Λ^\pm outlined in Eqs. (10)-(11) and the limiter matrices Ψ^\pm outlined in Eqs. (56)-(57) are substituted in Eq. (55), the flux at the interface corresponds exactly to the conventional Yee-Roe flux outlined in Eq. (46).

The reason why the Yee-Roe flux has been rewritten in Steger-Warming flux-vector splitting form is to obtain positivity-preserving conditions on the limiter matrix, which can be derived following the approach outlined in Ref. [12]. This would yield the following ranges on the limiter matrices (see Appendix A for a derivation):

$$-\left| \frac{2 [G_{i+1}^-]_r}{[G_{i+1}^-]_r - [G_i^-]_r} \right| < [\Psi_{i+1/2}^-]_{r,r} < \left| \frac{2 [G_{i+1}^-]_r}{[G_{i+1}^-]_r - [G_i^-]_r} \right| \quad (58)$$

$$-\left| \frac{2 [G_{i-1}^+]_r}{[G_{i-1}^+]_r - [G_i^+]_r} \right| < [\Psi_{i-1/2}^+]_{r,r} < \left| \frac{2 [G_{i-1}^+]_r}{[G_{i-1}^+]_r - [G_i^+]_r} \right| \quad (59)$$

We can then combine Eq. (57) and the positivity-preserving condition (58), and further combine Eq. (56) and the positivity-preserving condition (59) to obtain positivity-preserving expressions for the limiter matrices Ψ^- and Ψ^+ , respectively:

$$[\Psi_{i+1/2}^-]_{r,r} = \max \left(- \left| \frac{\xi [G_{i+1}^-]_r}{[\Delta G_{i+1/2}^-]_r} \right|, \min \left(\frac{[L_{i+1} L_{i+1/2}^{-1} (\Lambda_{i+1/2} - |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2}]_r}{2 [\Delta G_{i+1/2}^-]_r}, \left| \frac{\xi [G_{i+1}^-]_r}{[\Delta G_{i+1/2}^-]_r} \right| \right) \right) \quad (60)$$

And similarly, we can obtain a positivity-preserving range for the limiter matrix Ψ^+ :

$$[\Psi_{i+1/2}^+]_{r,r} = \max \left(- \left| \frac{\xi [G_i^+]_r}{[\Delta G_{i+1/2}^+]_r} \right|, \min \left(\frac{[L_i L_{i+1/2}^{-1} (\Lambda_{i+1/2} + |\Lambda|_{i+1/2}) \Phi_{i+1/2} M_{i+1/2}]_r}{2 [\Delta G_{i+1/2}^+]_r}, \left| \frac{\xi [G_i^+]_r}{[\Delta G_{i+1/2}^+]_r} \right| \right) \right) \quad (61)$$

with the limiter matrix Φ defined in Eq. (48) and the vector M defined in Eq. (47). The positivity-preserving expressions for the limiter matrices expressed in Eqs. (60) and (61) are used in conjunction with the flux at the interface, Eq. (55), which can be rewritten in the following computationally efficient form:

$$F_{i+1/2} = L_i^{-1} G_i^+ + L_{i+1}^{-1} G_{i+1}^- + \frac{1}{2} L_i^{-1} \Psi_{i+1/2}^+ \Delta G_{i+1/2}^+ - \frac{1}{2} L_{i+1}^{-1} \Psi_{i+1/2}^- \Delta G_{i+1/2}^- \quad (62)$$

with $G^\pm \equiv \Lambda^\pm L U$ and $\Delta G_{i+1/2}^\pm \equiv G_{i+1}^\pm - G_i^\pm$. For the stencil to be positivity-preserving, the positive and negative eigenvalues Λ^\pm (which are used to calculate the vectors G^\pm) must be determined either as specified in Eqs. (21)-(22) or as specified in Eqs. (44)-(45), and the user-defined constant ξ must be less than 2:

$$0 < \xi < 2 \quad (63)$$

Because of errors due to round-off when using double-precision variables within the computer code, it is necessary to fix ξ to 1.99 for most problems. Further, and rather interestingly, lowering ξ to within the range $0.5 < \xi < 1.0$ can help prevent the solution from diverging to aphysical states when using high time steps, and can also help prevent convergence hangs when solving steady-state problems. For these reasons, ξ is given a value of 0.5 for all the test cases here considered.

6. Positivity-Preserving Time Step

In the previous sections, it was found that by imposing the condition that the coefficients of the neighbor nodes within the discretization equation must be positive, some first-order and second-order flux functions could be obtained that are positivity-preserving. However, such is not sufficient to ensure that the discretization equation conforms to the rule of the positive coefficients and that the solution remains positivity-preserving. Indeed, while the flux functions proposed above ensure that the coefficients of the neighbor nodes are positive, they do not ensure that the coefficient of the central node is positive. This issue is here addressed by determining the conditions necessary to obtain a central node coefficient that is positive for both first-order and second-order flux functions. Following the same steps as in Ref. [12], it can be easily demonstrated that this leads to the following condition on the time step:

$$\Delta t < \begin{cases} \Delta x / \left([\Lambda_i^+]_{r,r} - [\Lambda_i^-]_{r,r} \right) \quad \forall r & \text{for a first-order flux function} \\ \Delta x / \left(2 [\Lambda_i^+]_{r,r} - 2 [\Lambda_i^-]_{r,r} \right) \quad \forall r & \text{for a second-order flux function} \end{cases} \quad (64)$$

It is emphasized that, when computing the positivity-preserving restrictions on the time step for flux difference splitting schemes, it is important to use the expressions for Λ^\pm outlined in Section 4 above (see Eqs. (21)-(22) and Eqs. (44)-(45)). Interestingly, in regions of uniform properties, it can be shown that the positive and negative eigenvalues outlined in Section 4 are equal to the usual definition (i.e. $\Lambda^\pm = \frac{1}{2}(\Lambda \pm |\Lambda|)$), and that the time step restrictions expressed in Eq. (64) correspond to the one obtained from the CFL condition for a first-order flux function and to half of the one obtained from the CFL condition for a second-order flux function. However, such is not the case in regions with non-uniform properties, where the positive and negative eigenvalues proposed herein do not collapse to the usual definition. Then, the time step restrictions derived here can differ significantly from those given by the CFL condition, with a tenfold discrepancy not being unusual in the vicinity of contact surfaces or shock waves.

7. Interface Averaging

For the schemes proposed herein, the Jacobian at the interface (and its corresponding eigenvectors and eigenvalues) are not determined through the Roe average but rather through a type of arithmetic averaging. It is here preferred not to use the Roe average because its use is found to yield unreasonably large Mach numbers within vacuums, and this eventually leads to difficulties to maintain positivity due to round-off errors. Indeed, because the temperature is obtained from the thermal energy, and because the thermal energy is obtained by subtracting the kinetic energy from the total energy, the temperature becomes increasingly tainted with round-off errors when the kinetic energy is large compared to the thermal energy (which occurs when the Mach number becomes large). This problem can be mostly avoided by replacing the Roe average by a type of arithmetic average at the interface. Specifically, the Jacobian matrix at the interface between cells is obtained from the density, velocities, and speed of sound as follows:

$$A_{i+1/2} = A(\rho_{i+1/2}, a_{i+1/2}, u_{i+1/2}) \quad (65)$$

where the density ρ and velocity u at the interface are obtained through an arithmetic mean, while the sound speed a at the interface is fixed to the maximum between the left and right sound speeds:

$$\rho_{i+1/2} = \frac{1}{2}(\rho_i + \rho_{i+1}) , \quad u_{i+1/2} = \frac{1}{2}(u_i + u_{i+1}) , \quad a_{i+1/2} = \max(a_i, a_{i+1}) \quad (66)$$

Fixing the interface sound speed in this manner as opposed to an arithmetic mean is found to improve, albeit slightly, the resolution within viscous layers. The type of averaging at the interface proposed herein is not an “arithmetic averaging” *per se* as it involves a max function. Nonetheless, for simplicity, we shall refer to the latter as “arithmetic averaging” to distinguish it from the Roe average.

It is noted that by determining the Jacobian at the interface as outlined above, one important property of the Roe scheme is lost: that is, a contact discontinuity can not be captured exactly within one cell. Numerical experiments indicate, however, that such is not a source of concern. For instance, when solving contact discontinuities propagating in time in a shocktube (i.e. the Riemann problem), little or no difference can be observed between the arithmetic and the Roe averaging procedures. Similarly, little discernible difference in solutions could be observed within shocks and expansion fans occurring in typical aerodynamic flowfields. The only instance that the use of Roe averaging is found to improve somewhat the resolution is upon solving high-Reynolds-number laminar boundary layers. Even then, the gains in resolution are limited to the flow region very close to the leading edge, and such does not impact significantly the overall skin friction. Furthermore, numerous simulations of flows of practical interest indicate that an arithmetic averaging procedure introduces minimal dissipation within turbulent boundary layers, hence leading to a small number of nodes needed to obtain a grid-converged solution (see on this point the grid convergence studies in Ref. [16] where the Roe scheme is used along with an arithmetic averaging procedure to solve supersonic turbulent flows).

Additionally to preventing the Mach number from reaching excessively high values within vacuums, the arithmetic average is advantaged over the Roe average by being straightforward to deploy to arbitrary systems of conservation laws. Indeed, when the governing equations are not limited to the mass, momentum, and total energy transport equations, but also include other equations (such as the transport of turbulence kinetic energy and vibrational energy for instance), it is not clear how the Roe average procedure should be modified to ensure that contact discontinuities can be captured within one cell. On the other hand, the type of arithmetic averaging outlined herein can be easily deployed to such systems of conservation laws, while limiting the dissipation within viscous layers to a satisfyingly low amount.

8. Test Cases

The performance of the proposed positivity-preserving variants of the Roe and Yee-Roe schemes compared to the original methods is now assessed through some numerical experiments in 1D, 2D, and 3D. Although the flux functions proposed herein were derived in 1D, they can be extended to multiple dimensions through dimensional splitting. That is, the flux derivative along each dimension is discretized through a one-dimensional stencil function of the convective flux in the respective dimension, along with its associated eigenvectors and eigenvalues. Because there exists an infinity of possible eigenvector sets applicable to the 2D-3D Euler equations, and because the obtained solution depends on the eigenvectors, it is important to use the same set of eigenvectors and eigenvalues as used herein to reproduce the results shown below. In this paper, the eigenvectors and eigenvalues are as specified in the Appendix of Ref. [12], because of their capability to capture waves with high-resolution even when utilized in conjunction with discretization stencils that enforce positivity-preservation. Unless otherwise indicated, the proposed method does not make use of an entropy correction. In “standard form” the flux is found from Eq. (62) with the Λ^\pm

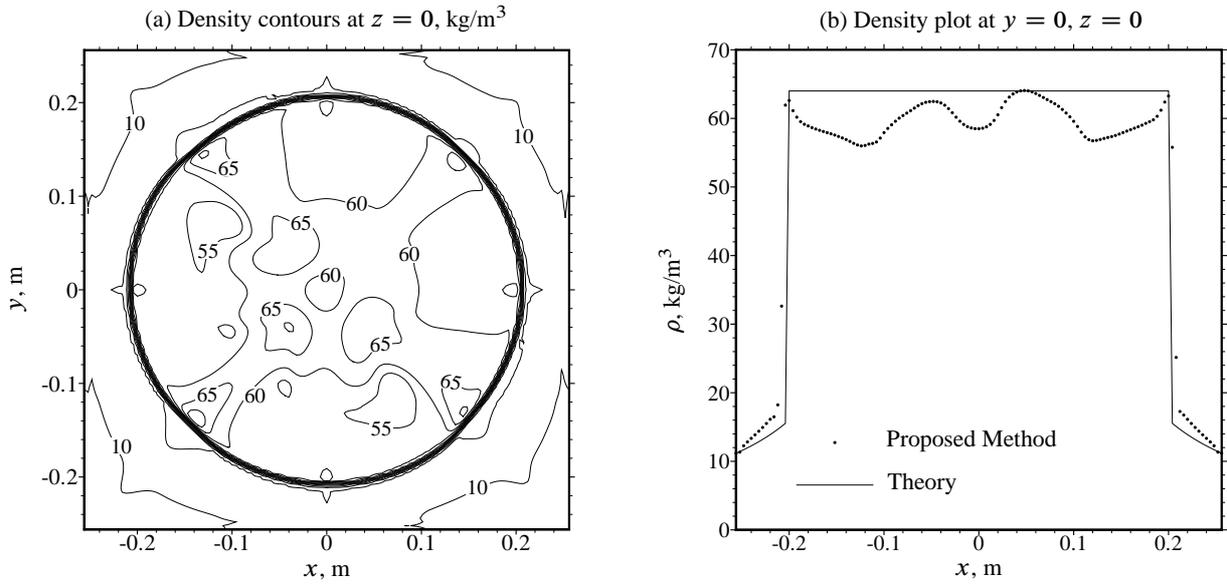


FIGURE 1. Density obtained with the proposed method at a time of 0.6 s for the 3D Noh test case; the entropy correction factor δ is set to 0.3.

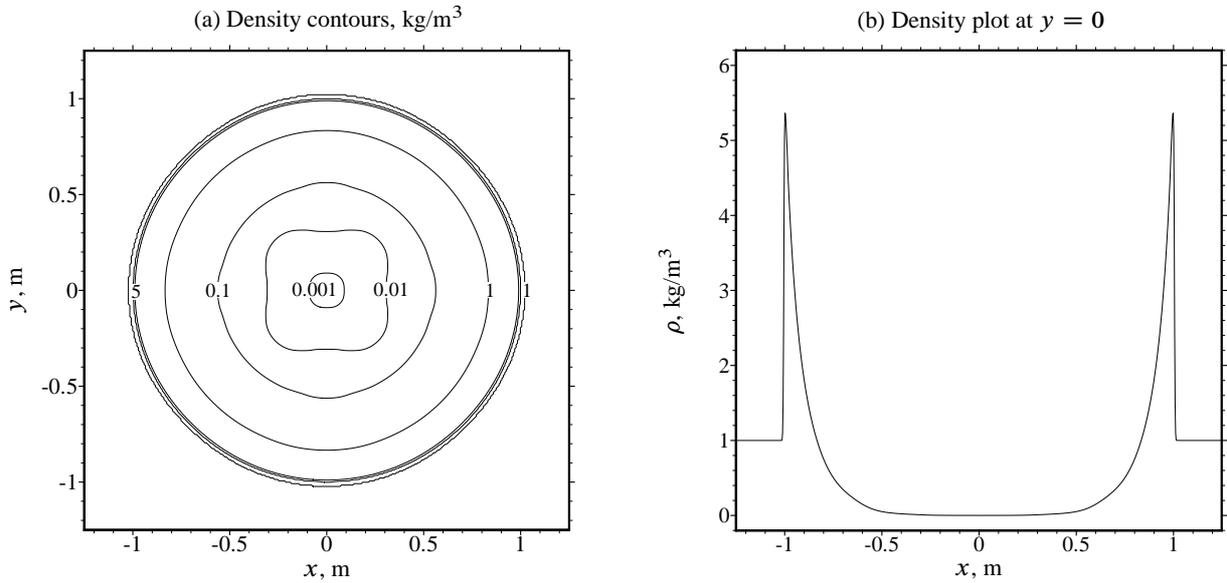


FIGURE 2. Density obtained with the proposed method at a time of 1 s for the 2D Sedov test case; the entropy correction factor δ is set to 0.3.

eigenvalues outlined in Eqs. (21) and (22). In “iterative form”, the flux is determined from Eq. (62) with the Λ^\pm eigenvalues outlined in Eqs. (44) and (45). Further, it is emphasized that the arithmetic averaging presented in the previous section is used only for the proposed stencils. For the original (non-positivity-preserving) Roe and Yee-Roe schemes, the Jacobian at the interface is rather obtained through Roe averaging.

This section is divided in two parts. First, we assess the capability of the proposed method to preserve the positivity of the density and pressure for several test cases that are known to be particularly stringent. This is then followed by another series of test cases to determine whether the positivity-preserving variant of the Roe solver proposed herein can capture as well as the original scheme some key flow features (such as viscous layers, rarefaction waves, shockwaves, etc).

8.1. Positivity-Preservation Capability Assessment

One well-known deficiency of the Roe scheme is the incapability to conserve the positivity of the density in the presence of strong rarefaction waves. Negative densities have been observed not only when a vacuum is created within the flow, but also for relatively mild rarefaction waves. The appearance of negative densities is exacerbated when the Roe flux function is turned second-order accurate through the Yee TVD limiting process. We now proceed to determine if the positivity-preserving variant of the Yee-Roe scheme that is proposed in this paper fixes this problem. For this purpose, we here consider two test cases that are known to present difficulties in maintaining positivity: the Sedov blast wave case and the Noh problem. Because both problems involve strong shock waves, it was deemed necessary to set the entropy correction factor δ to 0.3 to prevent aphysical phenomena from occurring.

First consider the particularly stringent 3D Noh problem, which consists of a gas with a specific heat ratio of 5/3 and a gas constant of 286 J/kgK with the following initial conditions: the inward radial velocity is set to 1 m/s, the pressure is set to as close to zero as possible, and the density is set to 1 kg/m³. Effectively, this entails initially a flow with an infinite Mach number directed towards the origin. To prevent a singularity, the initial pressure is not set to zero but rather to 10⁻⁶ Pa. The problem is here solved on a structured Cartesian mesh composed of 129³ equally-spaced nodes spanning a domain of (0.512 m)³. As the solution progresses in time, the boundary nodes are updated as follows: the pressure and the velocities remain fixed to the initial conditions while the density is updated according to the following exact solution [17, 18]:

$$\rho = \begin{cases} (1 + t/r)^2 & \text{for } r > t/3 \\ 64 & \text{for } r \leq t/3 \end{cases} \quad (67)$$

where r is in meters, t in seconds and the resulting ρ is in kg/m³. When solving this test case, the Yee-Roe flux function yields negative densities after a few iterations, even when using very small time steps several orders of magnitude below the CFL condition. On the other hand, as shown in Fig. 1a, the solution yielded by the proposed method is free of negative densities or temperatures as long as the time step is set to no more than one third the one yielded by the condition derived in Section 6.2 (it is noted that condition (64) is valid for a 1D system of conservation laws; in 2D, the time step needs to be reduced twofold, and in 3D, threefold). Further, as shown in Fig. 1b, the method proposed herein is seen to match reasonably well the theoretical solution despite the relatively low number of grid points used.

A second test case that can lead to some difficulties in preserving the positivity of the density and pressure is the Sedov blast wave case, which consists of a gas with a specific heat ratio of 1.4 and a gas constant of 286 J/kgK initially at rest with a density of 1 kg/m³ and a pressure of 10⁻⁶ Pa. The mesh is constructed of 1024² cells of equal width and height spanning a domain of (2.5 m)². For one cell at the origin, the pressure is set initially to 0.3917056 Pa · m² / A_{cell} with A_{cell} the area of the cell. The high pressure difference between the cell at the origin and its neighbors induces a strong cylindrical shockwave followed by a rarefaction wave, eventually leading to very low densities at the center of the domain. As shown in Fig. 2a, the solution yielded by the proposed method is free of negative densities or temperatures and displays good agreement with the exact solution: at a time of $t = 1$ s, the exact solution yields a shock located at a radius of 1 m and a density peak of 6 kg/m³ (see Refs. [19, 20]), which is corroborated by the results with the method proposed herein (see Fig. 2b).

Although not shown here for conciseness, the positivity-preserving capability of the proposed method has been further verified through the 2D Noh problem and the 3D Sedov problem, as well as through the stringent 1D, 2D, and 3D test cases outlined in Ref. [12]. The test cases outlined in the latter reference are such that the conventional Roe and Yee-Roe schemes fail to maintain the positivity of the density or the pressure, even for time steps well below the CFL condition. Such does not occur however using the stencils proposed herein: no negative densities or pressures were observed as long as the time step is fixed following the conditions presented in Section 6 above.

8.2. Accuracy Assessment for Flows of Interest

It is emphasized that the present approach achieves positivity-preservation by modifying both the first-order Roe terms and the second-order Yee terms as well as the type of averaging at the interface (i.e. arithmetic average instead of Roe average). It may be argued that these modifications result in stencils that do not retain the appealing features of the original schemes. For this purpose, we now proceed to determine through the simulation of some key problems if the proposed stencils do perform as well as the original Roe and Yee-Roe methods with respect to high-resolution capture of discontinuities, low dissipation within viscous layers, and ease of convergence to steady-state.

One appealing attribute of the Yee-Roe flux function is the capability to capture viscous layers while introducing a minimal amount of dissipation. It can be verified if the present approach retains this attribute through the simulation of a laminar

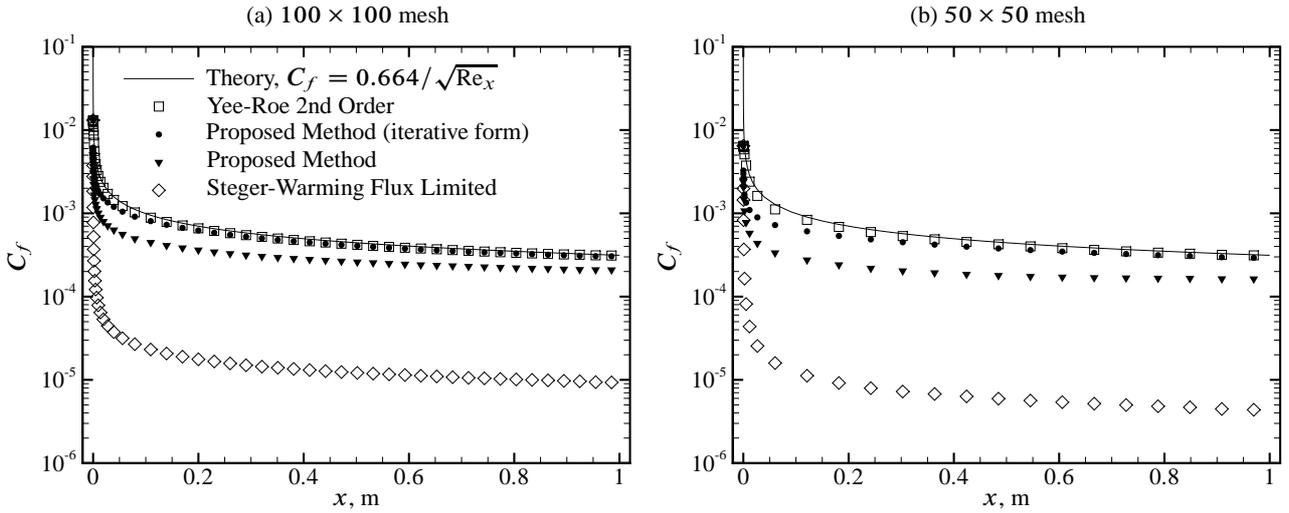


FIGURE 3. Comparison between the proposed method (using arithmetic averaging), the Yee-Roe second-order scheme (using Roe averaging), and the Steger-Warming flux limited scheme on the basis of skin friction coefficient at the wall for the boundary layer test case; one node in two is shown.

boundary layer over a flat plate. The simulation of high-Reynolds-number laminar boundary layers is especially difficult because the molecular diffusion taking place within laminar boundary layers is often of the same order of magnitude (or even less) than the aphysical dissipation introduced by the flux discretization scheme. In fact, several commonly-used flux functions (such as the Steger-Warming FVS method, or the Jameson second-fourth order artificial dissipation scheme, HLL, etc) introduce such a large amount of aphysical dissipation that an acceptable estimate for the skin friction can only be obtained if more than hundreds or even thousands of grid lines are clustered within the boundary layer.

To quantify the amount of aphysical dissipation introduced within viscous layers, consider air flowing over a flat plate with a Mach number of 2, a pressure of 0.1 bar, and a temperature of 300 K. The results are obtained using an orthogonal mesh which is constructed such that most of the nodes are distributed within the boundary layer (specifically: 60% of the gridlines are distributed, equally-spaced, within 2 mm of the wall). In Fig. 3, a comparison is offered between the proposed stencils and some conventional second-order Roe and Steger-Warming flux functions on the basis of the skin friction coefficient at the wall. The Yee-Roe scheme (with Roe averaging at the interface) can be seen to perform admirably well for this problem: even for the coarsest mesh considered, it yields a solution that is essentially grid-converged and that is in near-perfect agreement with the theoretical prediction (the small discrepancy is due to the analytical solution being inapplicable near the leading edge). On the other hand, the Steger-Warming scheme introduces excessive dissipation and yields a skin friction at the wall several orders of magnitude lower than the theoretical prediction, even for the finest mesh considered. Although not shown here, a grid convergence study indicates that, in order to match the resolution of the Yee-Roe flux function, the Steger-Warming method would require a mesh that is at least 100 times more refined near the wall. On the other hand, the present approach performs significantly better than the Steger-Warming scheme, by exhibiting a resolution approaching the one of the Yee-Roe method: on either coarse or fine meshes, the skin friction coefficient is within a few percent of the analytical solution over most of the flat plate (see Fig. 3). Several other simulations of laminar and turbulent flows over flat plates confirm that the present approach performs essentially as well as the original Yee-Roe method in capturing either boundary or shear layers, while being positivity-preserving and using an arithmetic average at the interface.

A second appealing attribute of the Yee-Roe flux function is the capability to capture discontinuities with “high-resolution”. High resolution here refers to the property of a method to capture with few nodes discontinuous or continuous waves while not introducing spurious non-physical oscillations. Roe-based flux functions perform very well in this regard compared to other methods. Indeed, compared to the Roe scheme, many other discretization stencils are either more dissipative (eg. HLL, Steger-Warming) or introduce more spurious oscillations (eg. Jameson 2nd-4th order artificial dissipation, AUSM) in the vicinity of contact surfaces or shockwaves. The high-resolution capability of the Yee-Roe method is due, in part, to the first-order Roe terms being monotonicity-preserving while introducing minimal dissipation in the vicinity of discontinuities, and, in part, to the second-order Yee terms being total variation diminishing (i.e., capability to maintain the monotonicity-preserving property of the underlying first-order scheme while reaching second-order accuracy in smooth flow regions). Because the present approach

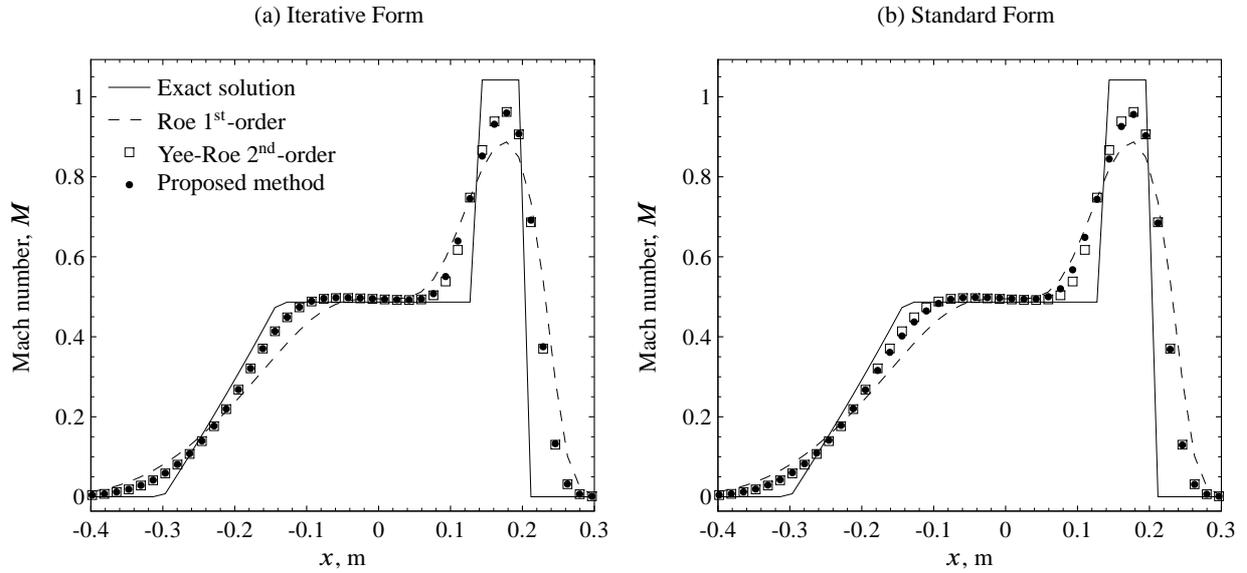


FIGURE 4. Comparison between the proposed method (using arithmetic averaging) and the Yee-Roe second-order scheme (using Roe averaging) for the Riemann problem test case at a time of 0.8 ms.

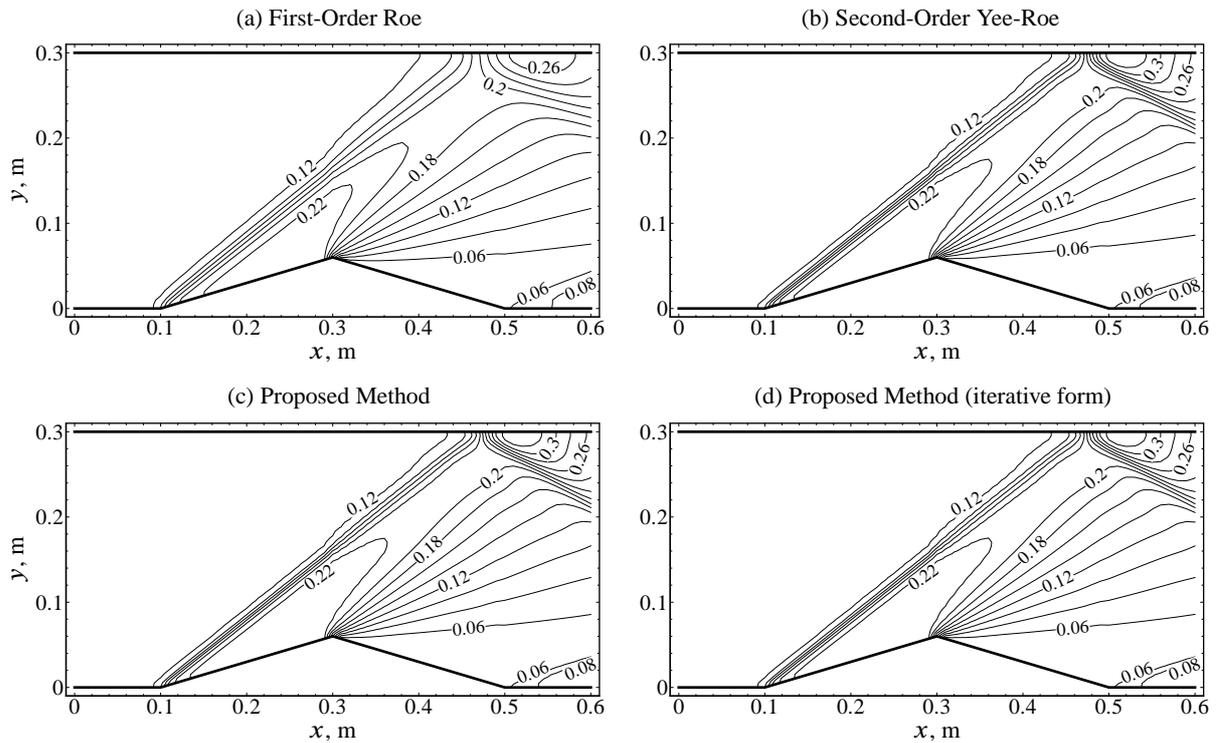


FIGURE 5. Comparison between the proposed method (using arithmetic averaging), and the Yee-Roe scheme (using Roe averaging) for the Mach 2.5 channel flow test case on the basis of the steady-state density contours (in kg/m^3) obtained using a 61×61 mesh.

modifies both the first-order Roe terms and the second-order Yee terms in order to attain positivity-preservation, it is important to quantify the amount of resolution lost due to these modifications. For this purpose, consider air with a density of 1 kg/m^3 initially at rest in a constant-area duct. The air is given initially a pressure of 1 bar for $x < 0.5 \text{ m}$ and a pressure of 0.1 bar for $x \geq 0.5 \text{ m}$, and the solution is advanced in time through an explicit Euler algorithm. As can be seen through the Mach number

profiles plotted in Fig. 4, the present approach yields a solution that is nearly identical to the one of the original Yee-Roe flux function: very little difference can be observed through the contact surface, and essentially no difference is apparent through the shockwave or the expansion fan. Although not shown here for conciseness, several other test cases in 1D, 2D, and 3D yield a similar conclusion, and confirm that the enforcement of positivity-preservation through the rule of the positive coefficients and the replacement of the Roe average by an arithmetic average do not diminish significantly the high-resolution capabilities of the Yee-Roe scheme.

A third appealing attribute of the Yee-Roe scheme is its capability to yield a converged solution for a wide variety of flows. A solution is considered "converged" when the discretized equations on all nodes are solved to an acceptable level of accuracy. Obtaining a converged solution is trivial when using an explicit Euler time marching algorithm to solve time-accurate cases, because an explicit Euler scheme solves the discretized equations on each node exactly. However, when solving time-accurate problems through an implicit dual-time stepping strategy or, alternately, when solving steady-state problems through a pseudotime relaxation procedure, there is no guarantee that the solution yielded by the flow solver is converged, even after a large number of iterations. In certain cases, obtaining a converged solution can be problematic because of convergence hangs originating from the limited second-order terms. Compared to alternative approaches (especially other TVD schemes), the Yee-Roe method fares well in this regard and offers good convergence behavior for various types of flows. To verify if the positivity-preserving variant of the Yee-Roe scheme proposed herein exhibits as good convergence characteristics as the original method, various types of steady-state problems have been investigated: in all cases, not only did the present approach converge as well as the original scheme, but little or no discernible differences could be found in the obtained solution. For instance, consider a steady inviscid supersonic flow over a triangle in a channel with a Mach number of 2.5, a temperature of 300 K, and a pressure of 0.1 bar. Through an iterative pseudotime-stepping relaxation process, the residual of the discretized equations on all nodes is minimized sufficiently that a converged solution is obtained. As can be seen through the steady-state density contours in Fig. 5, the present approach yields a solution that is essentially identical to the one of the Yee-Roe flux function with no apparent discrepancy within either the shocks or expansion fans.

9. Conclusions

A new positivity-preserving variant of the Roe flux difference splitting scheme is obtained by modifying the Roe flux function such that the associated discretization equation conforms to the rule of the positive coefficients. To satisfy the rule of the positive coefficients, all the coefficients within the discretization equation should have positive eigenvalues and have the same eigenvectors as those of the flux Jacobian. Because the modification does not alter the Roe wave speeds at the interface (the amount of wave speed lost by the right node is gained by the left node, and vice-versa), the modified stencils retain the favorable features of the original method such as high-resolution capture of discontinuities and of viscous layers. The positivity-preserving variant of the Roe scheme is extended to second-order accuracy through the Yee centered TVD limiters applied to the characteristic variables. To ensure that the second-order flux function is positivity-preserving, the Yee limiting process is altered such that it results in a discretization equation obeying the rule of the positive coefficients. Several test cases indicate that the extra amount of dissipation necessary to ensure positivity-preservation is typically negligible, and only becomes significant when the original method introduces negative densities and temperatures.

By imposing the rule of the positive coefficients on the discretization coefficient in which the time step appears, a positivity-preserving condition on the time step applicable to flux difference splitting schemes is derived. The so-obtained time step condition is shown to revert to the CFL condition in regions of uniform properties, but to depart from the latter in regions with appreciable property gradients. This is confirmed through various numerical experiments: near contact discontinuities or shock waves, the time step needed to ensure positivity preservation of flux difference splitting schemes can differ by an order of magnitude or more from the one obtained from the CFL condition.

In order to prevent the Mach number from reaching excessively high values in vacuums, it is found necessary to replace the Roe average by a type of arithmetic averaging at the interface. In so-doing, the scheme can not capture a contact discontinuity exactly within one cell. However, numerous test cases indicate that this is not a cause for concern: the replacement of Roe averaging by arithmetic averaging is found to have negligible impact on the solution except near the leading edge of high-Reynolds-number laminar boundary layers. Even in such regions, the dissipation introduced by the arithmetic average is minimal, and the shear stress and boundary layer height are close to those obtained with the Roe average. Other than preventing the Mach number from reaching too-high values within vacuums, arithmetic averaging is advantaged by being straightforward to extend to arbitrary systems of conservation laws.

Compared to previous positivity-preserving variants of the Roe scheme, the proposed method is noteworthy by being written in general matrix form. That is, the flux function depends solely on the vector of conserved variables, the convective flux vector,

and the convective flux Jacobian and its associated eigenvectors and eigenvalues. Therefore, the present method can solve not only the 1D Euler equations, but can be readily deployed to other systems of conservation laws such as the 2D-3D Euler equations in generalized coordinates (as done herein for several test cases). Positivity-preservation is guaranteed as long as the system of conservation laws adheres to the rule of the positive coefficients. Although the proof of the rule of the positive coefficients is limited to the perfect-gas Euler equations, some preliminary theoretical analysis and numerical experiments suggest that it may also be applicable to other fluid flow governing equations, including real gas effects, transport of multiple species densities, transport of turbulence kinetic energy, etc. Further study is nonetheless required to substantiate these claims, and to confirm whether the flux difference splitting schemes presented herein remain positivity-preserving for more intricate sets of governing equations.

A. Derivation of Positivity-Preserving Range on Limiter Function

The positivity-preserving limits on the limiter function are here derived for the Yee-Roe flux. This can be accomplished by first substituting the Yee-Roe flux function, Eq. (55), into the discrete equation, Eq. (12):

$$\begin{aligned} \frac{\Delta x}{\Delta t}(U_i^{n+1} - U_i) &= -L_i^{-1}\Lambda_i^+L_iU_i - L_{i+1}^{-1}\Lambda_{i+1}^-L_{i+1}U_{i+1} - \frac{1}{2}L_i^{-1}\Psi_{i+1/2}^+(G_{i+1}^+ - G_i^+) + \frac{1}{2}L_{i+1}^{-1}\Psi_{i+1/2}^-(G_{i+1}^- - G_i^-) \\ &+ L_{i-1}^{-1}\Lambda_{i-1}^+L_{i-1}U_{i-1} + L_i^{-1}\Lambda_i^-L_iU_i + \frac{1}{2}L_{i-1}^{-1}\Psi_{i-1/2}^+(G_i^+ - G_{i-1}^+) - \frac{1}{2}L_{i-1}^{-1}\Psi_{i-1/2}^-(G_i^- - G_{i-1}^-) \end{aligned} \quad (\text{A.1})$$

Then, following the approach outlined in Ref. [12], the latter can also be written as:

$$C_i^{n+1}U_i^{n+1} = C_iU_i + C_{i+1}U_{i+1} + C_{i-1}U_{i-1} \quad (\text{A.2})$$

In order for the latter to be equal to the former, the discretization coefficients must be defined as:

$$C_i^{n+1} \equiv \frac{\Delta x}{\Delta t}I \quad (\text{A.3})$$

$$C_iU_i \equiv \frac{\Delta x}{\Delta t}IU_i - L_i^{-1}G_i^+ - \frac{1}{2}L_i^{-1}\Psi_{i+1/2}^+(G_{i+1}^+ - G_i^+) + L_i^{-1}G_i^- - \frac{1}{2}L_i^{-1}\Psi_{i-1/2}^-(G_i^- - G_{i-1}^-) \quad (\text{A.4})$$

$$C_{i-1}U_{i-1} \equiv +L_{i-1}^{-1}G_{i-1}^+ + \frac{1}{2}L_{i-1}^{-1}\Psi_{i-1/2}^+(G_i^+ - G_{i-1}^+) \quad (\text{A.5})$$

$$C_{i+1}U_{i+1} \equiv -L_{i+1}^{-1}G_{i+1}^- + \frac{1}{2}L_{i+1}^{-1}\Psi_{i+1/2}^-(G_{i+1}^- - G_i^-) \quad (\text{A.6})$$

According to the rule of the positive coefficients, a discretization stencil is positivity-preserving if all the coefficients are positive, with a positive coefficient having all-positive eigenvalues and having the same eigenvectors as those of the flux Jacobian. It can be readily seen that the coefficient C_i^{n+1} satisfies these requirements. Further, the coefficient C_i becomes positive for a small-enough time step. However, such is not the case for the other two coefficients. Let us now determine the conditions on the limiters matrices Ψ^\pm that ensure that the coefficients C_{i-1} and C_{i+1} are positive.

For the coefficient C_{i+1} to be positive, it must have the same eigenvectors as those of the flux Jacobian and have all-positive eigenvalues. Then, it follows that Eq. (A.6) must correspond to:

$$L_{i+1}^{-1}D_{i+1}^+L_{i+1}U_{i+1} = -L_{i+1}^{-1}G_{i+1}^- + \frac{1}{2}L_{i+1}^{-1}\Psi_{i+1/2}^-(G_{i+1}^- - G_i^-) \quad (\text{A.7})$$

with D_{i+1}^+ denoting a diagonal matrix with positive diagonal elements. Multiply all terms by $2L_{i+1}$ and express the terms in tensor form:

$$2[D_{i+1}^+]_{r,r}[L_{i+1}U_{i+1}]_r = -2[G_{i+1}^-]_r + [\Psi_{i+1/2}^-]_{r,r}[G_{i+1}^-]_r - [\Psi_{i+1/2}^-]_{r,r}[G_i^-]_r \quad (\text{A.8})$$

Multiply all terms by $[\Lambda_{i+1}^-]_{r,r}$ and divide through by $[G_{i+1}^-]_r$:

$$2[D_{i+1}^+]_{r,r} = -2[\Lambda_{i+1}^-]_{r,r} + [\Lambda_{i+1}^-]_{r,r}[\Psi_{i+1/2}^-]_{r,r} - [\Lambda_{i+1}^-]_{r,r}[\Psi_{i+1/2}^-]_{r,r}\frac{[G_i^-]_r}{[G_{i+1}^-]_r} \quad (\text{A.9})$$

For the stencil to be positivity-preserving, all diagonal elements of the matrix $[D_{i+1}^+]_{r,r}$ must be positive. Then it follows that the following condition must hold:

$$-2[\Lambda_{i+1}^-]_{r,r} + [\Lambda_{i+1}^-]_{r,r}[\Psi_{i+1/2}^-]_{r,r} - [\Lambda_{i+1}^-]_{r,r}[\Psi_{i+1/2}^-]_{r,r}\frac{[G_i^-]_r}{[G_{i+1}^-]_r} > 0 \quad (\text{A.10})$$

Divide through by $[\Lambda_{i+1}^-]_{r,r}$ and rearrange:

$$[\Psi_{i+1/2}^-]_{r,r} \frac{[G_{i+1}^-]_r - [G_i^-]_r}{[G_{i+1}^-]_r} < 2 \quad (\text{A.11})$$

To obtain the latter, it is noted that we divided by $[\Lambda_{i+1}^-]_{r,r}$, which is either negative or zero when using the positivity-preserving variant outlined in Eq. (21) or in Eq. (45). A division by zero could be avoided by redefining the negative eigenvalues such that they never exceed a negative number of small magnitude. However, such is not deemed necessary because the potential division by zero only appears in an interim equation and not in the final flux function.

If the LHS of (A.11) is negative, the condition is guaranteed to be satisfied. Therefore, we can take the absolute value of the terms on the LHS without losing generality:

$$\left| [\Psi_{i+1/2}^-]_{r,r} \frac{[G_{i+1}^-]_r - [G_i^-]_r}{[G_{i+1}^-]_r} \right| < 2 \quad (\text{A.12})$$

Further, it can be easily shown that condition (A.12) yields the following bounds on the limiter function Ψ^- :

$$-\left| \frac{2[G_{i+1}^-]_r}{[G_{i+1}^-]_r - [G_i^-]_r} \right| < [\Psi_{i+1/2}^-]_{r,r} < \left| \frac{2[G_{i+1}^-]_r}{[G_{i+1}^-]_r - [G_i^-]_r} \right| \quad (\text{A.13})$$

Having found the conditions on the limiter function Ψ^- that yield a positivity-preserving stencil, proceed to determine the conditions on the limiter function Ψ^+ . This can be accomplished starting from Eq. (A.5) and following the same steps as above. The following is thus obtained:

$$\left| [\Psi_{i-1/2}^+]_{r,r} \frac{[G_{i-1}^+]_r - [G_i^+]_r}{[G_{i-1}^+]_r} \right| < 2 \quad (\text{A.14})$$

The latter yields the following bounds on the limiter function Ψ^+ :

$$-\left| \frac{2[G_{i-1}^+]_r}{[G_{i-1}^+]_r - [G_i^+]_r} \right| < [\Psi_{i-1/2}^+]_{r,r} < \left| \frac{2[G_{i-1}^+]_r}{[G_{i-1}^+]_r - [G_i^+]_r} \right| \quad (\text{A.15})$$

Acknowledgment

This work was supported for two years by a Pusan National University Research Grant.

References

- [1] ROE, P. L., "Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes," *Journal of Computational Physics*, Vol. 43, 1981, pp. 357–372.
- [2] ROE, P. L., "Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes," *Journal of Computational Physics*, Vol. 135, 1997, pp. 250–258.
- [3] GODUNOV, S. K., "A Finite Difference Method for the Numerical Computation of Discontinuous Solutions of the Equations of Fluid Dynamics," *Mathematics Sbornik*, Vol. 47, No. 3, 1959, pp. 271–306.
- [4] TORO, E. F., SPRUCE, M., AND SPEARES, W., "Restoration of the Contact Surface in the HLL-Riemann Solver," *Shock Waves*, Vol. 4, 1994, pp. 25–34.
- [5] EDWARDS, J. R. AND LIOU, M.-S., "Low-Diffusion Flux-Splitting Methods for Flows at All Speeds," *AIAA Journal*, Vol. 36, 1998, pp. 1610–1617.
- [6] STEGER, J. L. AND WARMING, R. F., "Flux Vector Splitting of the Inviscid Gasdynamics Equations with Application to Finite Difference Methods," *Journal of Computational Physics*, Vol. 40, 1981, pp. 263–293.
- [7] HARTEN, A., LAX, P. D., AND VAN LEER, B., "On Upstream Differencing and Godunov-type Schemes for Hyperbolic Conservation Laws," *SIAM Review*, Vol. 25, No. 1, 1983, pp. 35–61.
- [8] EINFELDT, B., MUNZ, C. D., ROE, P. L., AND SJOGREEN, B., "On Godunov-type Methods near Low Densities," *Journal of Computational Physics*, Vol. 92, 1991, pp. 273–295.
- [9] DUBROCA, B., "Positively Conservative Roe's Matrix for Euler Equations," *Lecture Notes in Physics*, Vol. 515, 1998, pp. 272–277, Proceedings of the 16th International Conference of Numerical Methods for Fluid Dynamics.
- [10] DUBROCA, B., "Solveur de Roe Positivement Conservatif," *Comptes Rendus de l'Académie des Sciences – Series I – Mathematics*, Vol. 329, No. 9, 1999, pp. 827–832.
- [11] PARENT, B., "Positivity-Preserving Flux-Limited Method for Compressible Fluid Flow," *Computers & Fluids*, Vol. 44, No. 1, 2011, pp. 238–247.

- [12] PARENT, B., "Positivity-Preserving High-Resolution Schemes for Systems of Conservation Laws," *Journal of Computational Physics*, Vol. 231, No. 1, 2012, pp. 173–189.
- [13] YEE, H. C., KLOPFER, G. H., AND MONTAGNÉ, J.-L., "High-Resolution Shock-Capturing Schemes for Inviscid and Viscous Hypersonic Flows," *Journal of Computational Physics*, Vol. 88, 1990, pp. 31–61.
- [14] YEE, H. C., "Construction of Explicit and Implicit Symmetric TVD Schemes and Their Applications," *Journal of Computational Physics*, Vol. 68, 1987, pp. 151–179.
- [15] YEE, H. C., "A Class of High-Resolution Explicit and Implicit Shock Capturing Methods," TM 101088, NASA, 1989.
- [16] PARENT, B. AND SISLIAN, J. P., "Validation of Wilcox $k\omega$ Model for Flows Characteristic to Hypersonic Airbreathing Propulsion," *AIAA Journal*, Vol. 42, No. 2, 2004, pp. 261–270.
- [17] NOH, W. F., "Errors for Calculations of Strong Shocks Using an Artificial Viscosity and an Artificial Heat Flux," *Journal of Computational Physics*, Vol. 72, 1987, pp. 78–120.
- [18] JOHNSEN, E., LARSSON, J., BHAGATWALA, A. V., CABOT, W. H., MOIN, P., OLSON, B. J., RAWAT, P. S., SHANKAR, S. K., SJOGREEN, B., YEE, H. C., ZHONG, X., AND LELE, S. K., "Assessment of High-Resolution Methods for Numerical Simulations of Compressible Turbulence with Shock Waves," *Journal of Computational Physics*, Vol. 229, 2010, pp. 1213–1237.
- [19] KAMM, J. R. AND TIMMES, F. X., "On Efficient Generation of Numerically Robust Sedov Solutions," LA-UR 07-2849, Los Alamos National Laboratory, 2007.
- [20] MAIRE, P.-H., "A High-Order Cell-Centered Lagrangian Scheme for Two-Dimensional Compressible Fluid Flows on Unstructured Meshes," *Journal of Computational Physics*, Vol. 228, 2009, pp. 2391–2425.