

# On the implementation of flux limiters in algebraic frameworks <sup>☆</sup>

Nicolás Valle <sup>a</sup>, Xavier Álvarez-Farré <sup>a</sup>, Andrey Gorobets <sup>b</sup>, Jesús Castro <sup>a,\*</sup>, Assensi Oliva <sup>a</sup>, F. Xavier Trias <sup>a</sup>



<sup>a</sup> Heat and Mass Transfer Technological Center, Technical University of Catalonia, Carrer Colom 11, 08222 Terrassa (Barcelona), Spain

<sup>b</sup> Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya Sq. 4, 125047 Moscow, Russia

## ARTICLE INFO

### Article history:

Received 1 June 2021

Received in revised form 5 October 2021

Accepted 9 November 2021

Available online 16 November 2021

### Keywords:

Flux limiter

Parallel CFD

Heterogeneous computing

Portability

Mimetic

## ABSTRACT

The use of flux limiters is widespread within the scientific computing community to capture shock discontinuities and are of paramount importance for the temporal integration of high-speed aerodynamics, multiphase flows and hyperbolic equations in general.

Meanwhile, the breakthrough of new computing architectures and the hybridization of supercomputer systems pose a huge portability challenge, particularly for legacy codes, since the computing subroutines that form the algorithms, the so-called kernels, must be adapted to various complex parallel programming paradigms. From this perspective, the development of innovative implementations relying on a minimalist set of kernels simplifies the deployment of scientific computing software on state-of-the-art supercomputers, while it requires the reformulation of algorithms, such as the aforementioned flux limiters.

Equipped with basic algebraic topology and graph theory underlying the classical mesh concept, a new flux limiter formulation is presented based on the adoption of algebraic data structures and kernels. As a result, traditional flux limiters are cast into a stream of only two types of computing kernels: sparse matrix-vector multiplication and generalized pointwise binary operators. The newly proposed formulation eases the deployment of such a numerical technique in massively parallel, potentially hybrid, computing systems and is demonstrated for a canonical advection problem.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The evolution in hardware technologies enables scientific computing to advance incessantly and reach further aims. Nowadays, the use of high-performance computing (HPC) systems is rather common in the solution of both industrial and academic scale problems. However, many algorithms employed in scientific computing have a very low arithmetic intensity (AI), which is the ratio of computing work in floating-point operations (FLOP) to memory traffic in bytes, hence numerical simulation codes are usually memory-bounded, making processors suffer from serious data starvation [1,2]. To top it off, the calculations often result in irregular, non-coalescing memory access patterns, reducing the memory access efficiency. Ironically, the memory bandwidth of computing hardware grows much slower than its peak performance, aggravating the problem. All of this motivates the introduction of new parallel architectures with faster and more complex memory configurations into HPC systems.

To take advantage of the increasing variety of computing architectures and the hybridization of HPC systems, the computing subroutines that form the algorithms, the so-called kernels, must be adapted to complex paradigms such as distributed-memory (DM) and shared-memory (SM) multiple instruction, multiple data (MIMD) parallelism, and stream processing (SP). It also encourages the demand for portable and sustainable implementations of scientific simulation codes [3]. While portability is an intangible characteristic of software, it may be easy for a developer to have an idea of how difficult it is to rewrite, debug and verify a specific code on its adaptation to a new architecture. On the other hand, sustainability refers to developing reusable and resilient codes. The way a code is conceived at its inception enormously determines the degree to which both properties can be attained.

Traditionally, the development of scientific computing software is based on calculations in iterative stencil loops (ISL) over a discretized geometry—the mesh. This implementation approach is referred to as *stencil-based*. Despite being intuitive and versatile, the interdependency between algorithms and their computational implementations in stencil applications usually results in a large number of subroutines and introduces an inevitable complexity when it comes to portability and sustainability [4].

<sup>☆</sup> The review of this paper was arranged by Prof. N.S. Scott.

\* Corresponding author.

E-mail address: [jesus.castro@upc.edu](mailto:jesus.castro@upc.edu) (J. Castro).

Regarding portability, the complexity of stencil applications motivates the adoption of conservative strategies, which consist of porting (rewriting) the most time-consuming part of an existing code, or even the entire code, to a new architecture but minimizing the structural modifications. In other words, it leads to a partial or complete reimplementations of an existing code. These strategies were common during the rise of general-purpose computing on graphics processing units (GPUs) because they allow for direct comparison studies of both numerical and performance results versus the legacy versions. Well-known commercial computational fluid dynamics (CFD) codes and open-source platforms offer GPU extensions for solvers of systems of linear algebra equations (SLAE), which represent a significant part of the overall computing time. This can provide substantial acceleration with compactly localized changes in the code. Such an example can be found in [5], where the authors coupled a GPU-accelerated library for solving large sparse SLAE with the OpenFOAM platform and demonstrated performance on up to 128 nodes of a GPU-based cluster. The use of GPUs in scientific computing is nowadays rather mature, and there are many successful examples in the literature [6–9]. For instance, the early GPU implementations in [10], extended in [11], proved to be two orders of magnitude faster than its central processing unit (CPU) counterpart. Moreover, the solution of two-phase flows on multi-GPU systems [12] was not only faster but also more energy-efficient. An example of a GPU porting of an open-source Navier-Stokes solver, the AFid code, is found in [13]. Further examples of multi-GPU simulations of supersonic and hypersonic flows can be found in [14]. One of the most impressive GPU-based simulations is found in [15], after [16], on the solution of turbulent flows, reporting a sustained performance of 13.7PFLOPS.

Regarding sustainability, the implementation of new physics or numerical methods in a stencil-based framework, or its specialization for different mesh types, usually requires the design of new computing subroutines and data structures. This represents the main drawback of such an approach because the effort is not necessarily accumulative and thus reduces the software's sustainability. To address this, some authors propose domain-specific tools to generalize the stencil computations for specific fields. For instance, a framework that automatically translates stencil functions written in C++ to both CPU and GPU codes is proposed in [17]. However, these generalizations are still heavily restricted by the shape of the stencil they target.

An alternative to stencil implementations is to break the aforementioned interdependency between algorithm and implementation so that the calculations are cast into a minimalist set of kernels. In other words, the idea is to use the classical ISL just for building data and leave the calculations to a reduced set of basic operations; thus, legacy codes may be maintained indefinitely as preprocessing tools, and the calculation engines become easy to port and optimize.

By casting discrete operators and mesh functions into sparse matrices and vectors, it has been shown that nearly 90% of the calculations in a typical CFD algorithm for the direct numerical simulation (DNS) and large eddy simulation (LES) of incompressible turbulent flows boil down to the following basic linear algebra subroutines: sparse matrix-vector product ( $S_{PMV}$ ), linear combination of vectors ( $axpy$ ) and dot product ( $dot$ ) [18]. Moreover, after the generalizations detailed in Section 3.2 this value will be raised to 100%. Hereinafter, we refer to this implementation based on algebraic subroutines as *algebraic* or *algebra-based*. In this algebraic approach, the kernel code shrinks to dozens of lines; the portability becomes natural, and maintaining OpenMP, OpenCL, or CUDA implementations takes minor effort. Besides, standard libraries optimized for particular architectures (e.g., cuSPARSE [19], cISPARSE [20]) can be easily linked in addition to specialized in-house implementations. A similar approach is found in PyFR [16],

where the majority of operations are cast in terms of matrix-matrix multiplications linking with appropriate BLAS libraries. In the context of the DNS, the preconditioned conjugate gradient (CG) method following such an algebraic approach was implemented in [21], and its potential was exploited in [22] to perform petascale CFD simulations.

In our previous work, we proposed a heterogeneous implementation of an algebraic framework [23], the hierarchical parallel code for high-performance computing (HPC<sup>2</sup>), as a portable solution with many potential applications in the fields of computational physics and mathematics. Later, the minimalist design allowed us to easily optimize our framework for HPC systems with non-uniform memory access (NUMA) configurations [24], just by changing the definition of the three kernels and the initialization of vectors and matrices.

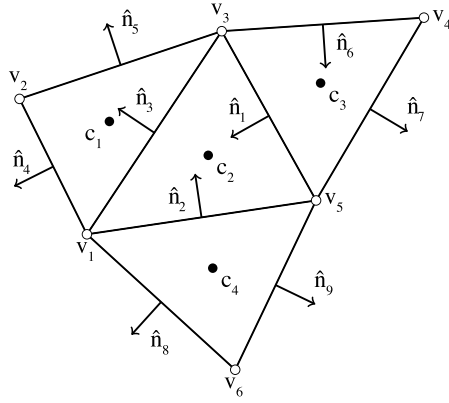
While linear schemes fit naturally in the proposed algebraic approach, the implementation of non-linear schemes is not evident. However, as a consequence of Godunov's theorem [25,26], those are required for the treatment of shock discontinuities to avoid unstable discretizations and the onset of wiggles. Such discontinuities are present in many industrial applications and appear in both compressible and multiphase flows, representing a classical problem in numerical analysis since, at least, the 1960's. The construction of stable, second-order (and higher) discretizations, then, requires the adoption of non-linear schemes to exhibit a total variation diminishing (TVD) [27] behavior. Among them, flux limiters are a mature and robust method, which has been adopted in a diversity of applications. Sweby [28] generalized several limiters and stated the conditions for second-order TVD schemes in a one-dimensional homogeneous mesh in its well-known *Sweby diagram*. Despite the known inconsistencies that arise when departing from the one-dimensional homogeneous case [29,30] these techniques have been ported to non-homogeneous Cartesian [30] and unstructured [31] meshes as well. Advances in this field have also been exploited by the multiphase flow community, particularly for the advection of the marker function [32,33].

Both the analysis and the implementation of flux limiters are typically performed from the aforementioned stencil-based perspective. However, the growing interest of the community in mimetic methods [34] unveils an alternative to the implementation of flux limiters. Mimetic methods construct discrete operators directly from the inherent incidence matrices that define the mesh. Adopting such an approach presents an important advantage from both theoretical and practical points of view. On the one hand, this allows for a flawless discrete mimicking of the continuum operators, facilitating the exact conservation of important secondary properties, such as energy [35,36], among others. On the other hand, this reduces the implementation to the right combination of a reduced set of algebraic subroutines. Therefore, the present work is devoted to the formulation of flux limiter schemes and their implementation into algebraic frameworks. The proposed implementation will be analyzed from a computational point of view and compared with the classical stencil counterpart, and the benefits of each option will be discussed in depth.

The rest of the paper is organized as follows. In section 2 a review of basic concepts of graph theory is briefly summarized in order to provide some context. Section 3 develops a generalization of flux limiters from an algebraic perspective and introduces the matrix-based calculation of the gradient ratio. Section 4 highlights the capabilities of the method on a well-known three-dimensional deformation benchmark. Finally, the discussion and conclusions are stated in Sections 5 and 6.

## 2. Algebraic topology

By using concepts from algebraic topology, mimetic methods preserve the inherent structure of the space, leading to stable and



**Fig. 1.** 2D mesh composed of cells  $c_i$ , which are bounded by faces  $f_j$  oriented in the direction  $\hat{n}_j$ . Faces, which collapse to the same entities as edges, are bounded by the set of vertices  $v_k$ .

robust discretizations [37,34]. However, the development of such techniques is out of the scope of this paper, where we rather focus on exploiting the relationships between the different entities of the mesh for the construction of flux limiters. The interested reader is referred to [38] and references therein.

Given whatever space of interest  $\Omega$ , we can equip it with a partition of unity, namely a mesh  $M$ , by bounding the group of cells,  $C$ , with faces,  $F$ ; those with the set of edges,  $E$ , and finally those with the set of vertices,  $V$ . In this sequence, groups are related to the next element of the sequence by means of the boundary operator  $\partial$ . This is known as a chain complex [37,39]. A two-dimensional example can be seen in Fig. 1, where faces and edges collapse into the same entity.

In every space  $Q$ , discrete variables are arranged in arrays such as  $\theta_q \in \mathbb{R}^{|Q|}$ . The relationship between the bounding elements of a geometric entity can be cast in oriented incidence matrices,  $E_{E \rightarrow V}$ ,  $E_{F \rightarrow E}$  and  $E_{C \rightarrow F}$ , corresponding to edge-to-vertex, face-to-edge and cell-to-face, respectively. The corresponding incidence matrices for the mesh depicted in Fig. 1 read:

$$E_{F \rightarrow V} = \begin{pmatrix} 0 & +1 & +1 & +1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & +1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 \\ +1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & +1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & -1 \end{pmatrix}, \quad (1)$$

$$E_{C \rightarrow F} = \begin{pmatrix} 0 & -1 & +1 & 0 \\ 0 & -1 & 0 & +1 \\ -1 & +1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & +1 \end{pmatrix}. \quad (2)$$

Conversely, we can define  $E_{F \rightarrow C}$ ,  $E_{E \rightarrow F}$  and  $E_{V \rightarrow E}$ , for the face-to-cell, edge-to-face and vertex-to-edge incidence matrix. Such converse incidence matrices are obtained by transposition:

$$\begin{aligned} E_{F \rightarrow C} &= E_{C \rightarrow F}^T, \\ E_{E \rightarrow F} &= E_{F \rightarrow E}^T, \\ E_{V \rightarrow E} &= E_{E \rightarrow V}^T. \end{aligned}$$

Incidence matrices represent the boundary operator between one element of the chain and the next one. Following the exam-

ple of Fig. 1,  $E_{F \rightarrow C}$  and  $E_{V \rightarrow F}$  provide with the orientation of the boundary faces  $f_j$  for every cell  $c_i$  and the boundary vertices  $v_k$  for every face  $f_j$ . Incidence matrices also play an essential role in preserving properties of the discrete space. In particular, they form an exact sequence. Exact sequences are those such that the application of the boundary operator twice results in 0. This can be verified by checking  $E_{F \rightarrow V} E_{C \rightarrow F} = 0$ . This property is shared by its continuum counterpart, the de Rham cohomology [34], which is the ultimate responsible of the following vector calculus identities [37]:

$$\nabla \times \nabla \equiv 0, \quad (3)$$

$$\nabla \cdot \nabla \times \equiv 0. \quad (4)$$

These are powerful identities that mimetic methods preserve by construction. For an extended review of the relationship between the continuum and the discrete counterparts, the reader is referred to [37,34] and references therein.

In addition to provide a suitable platform for the construction of appropriate mimetic methods, the relations contained in incidence matrices can be studied from a graph theory perspective.

A straightforward use of incidence matrices allows to compute differences across faces. The fact that differences lie in a different space (faces) than variables (cells) is an inherent property of such an approach:

$$\Delta \theta_c = E_{C \rightarrow F} \theta_c. \quad (5)$$

Particularly useful is the construction of undirected incidence matrices ( $B_{Q \rightarrow S}$ ), which are built by taking the absolute value of the elements of the directed ones ( $E_{Q \rightarrow S}$ ). Considering the index notation between a generic space  $Q$  (e.g., cells, faces) and its boundary  $S$  (e.g., faces, edges), we could proceed as follows:

$$[B_{Q \rightarrow S}]_{sq} = b_{sq} = |e_{sq}|, \quad (6)$$

where  $e_{sq} = [E_{Q \rightarrow S}]_{sq}$ .

Similarly, one can proceed to compute the degree matrix of the graph, which accounts for the number of connections that an entity has (e.g., the number of cells in contact with a face). Degree matrices are always diagonal and the value of the diagonal elements is obtained as follows:

$$W_{QQ} = \text{diag}(B_{S \rightarrow Q} \mathbb{1}_S). \quad (7)$$

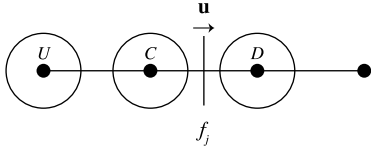
In particular, undirected incidence matrices can be used to construct suitable shift operators [40]:

$$\Pi_{C \rightarrow F} = W_{FF}^{-1} B_{C \rightarrow F}. \quad (8)$$

This provides with a simple face-centered interpolation, weighted with the number of adjacent faces. Note that by taking this approach, boundaries are inherently included from the graph information.

The use of such  $\Pi_{C \rightarrow F}$  is restricted to scalar fields. However, following [40], this can be readily extended to vector fields as follows. First, the discretization of a continuum vector field,  $\vec{u} = (u_1, \dots, u_d)^T$ , is arranged in a single column vector,  $\mathbf{u}_c \in \mathbb{R}^{d|C| \times 1} = (\mathbf{u}_1, \dots, \mathbf{u}_d)^T$ , where  $\mathbf{u}_i = ((u_i)_1, (u_i)_2, \dots, (u_i)_{|C|})^T$  are the vectors containing the components corresponding to the  $i$ th spatial direction. Note that  $d$  is the number of spatial dimensions of the problem. Next, the interpolator can be extended component-wise by applying the Kronecker product with the identity matrix of size  $\mathbb{R}^{d \times d}$ . The final ensemble is as follows:

$$\Gamma_{C \rightarrow F} = \mathbb{I}_d \otimes \Pi_{C \rightarrow F}. \quad (9)$$



**Fig. 2.** Classical stencil for the computation of the gradient ratio at face  $f_j$ .  $U$ ,  $C$  and  $D$  correspond to the upstream, centered and downstream nodes.

Similarly, normal vectors can be cast into a  $\mathbb{R}^{|F| \times d|F|}$  matrix by arranging  $d$  diagonal matrices, corresponding to every component of the face vector, next to each other as  $N_F = (N_1 | \dots | N_d)$  [40]. The two-dimensional  $N_F$  matrix corresponding to the mesh depicted in Fig. 1 reads:

$$N_F = \begin{pmatrix} n_{1x} & 0 & \dots & 0 & n_{1y} & 0 & \dots & 0 \\ 0 & n_{2x} & \dots & 0 & 0 & n_{2y} & \dots & \vdots \\ \vdots & 0 & \ddots & \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & n_{gx} & 0 & 0 & \dots & n_{gy} \end{pmatrix}. \quad (10)$$

In such a way, it is straightforward to either project a discrete vector as  $N_F \mathbf{x}_f$ , or to vectorize a scalar quantity as  $N_F^T \mathbf{s}_f$ , provided that both are stored at the faces. An accurate discussion about the construction of this matrix can be found in [40].

Other basic matrices derived from the graph are the graph Laplacian ( $L_{CC}$ ) and the adjacency matrix ( $A_{CC}$ ):

$$L_{CC} = B_{F \rightarrow C} B_{C \rightarrow F}, \quad (11)$$

$$A_{CC} = W_{CC} - L_{CC}. \quad (12)$$

Both are constructed based on the incidence matrices and provide information about the propagation of information along the graph. They are constructed by connecting cells to its neighbors through its bounding faces.

In summary, the constructor of such operators provides with tools able to relate different elements of the graph between each others. Equipped with such basic concepts, the development of higher level operators can proceed as in the following section.

### 3. Flux limiters

The solution of hyperbolic problems in finite volume methods when sharp discontinuities are present requires the use of high resolution schemes in order to attain second order approximations. In turn, the construction of such schemes is reduced to the appropriate reconstruction of the flux at the faces. Prone to introduce numerical instabilities, such a reconstruction requires an appropriate flux reconstruction strategy in order to guarantee TVD behavior (*i.e.*, such that no new minima or maxima are introduced). This is attained by limiting the flux at cell's boundaries by means of a flux limiter function.

Typically, flux limited schemes are stated in the following form [28]:

$$\theta_f = \theta_C + \Psi(r) \left( \frac{\theta_D - \theta_C}{2} \right), \quad (13)$$

where  $\theta_C$  and  $\theta_D$  stand for the centered and downwind values of  $\theta$  according to the velocity field  $u$  and  $\Psi(r)$  is the flux limiter function. Fig. 2 depicts this situation.

From a physical point of view, this is equivalent to the introduction of some sort of artificial diffusion which stabilizes the method at the expense of smearing out its profile. This can be easily seen by rewriting the classical stencil-based formulation stated in equation (13) into:

$$\theta_f = \frac{\theta_C + \theta_D}{2} + \frac{\Psi(r) - 1}{2} (\theta_D - \theta_C), \quad (14)$$

where  $(\Psi(r) - 1)/2$  stands for the artificial diffusion added to a classical symmetry-preserving scheme.

The limiting approach has been used by several authors [41,26] who over the years developed several discontinuity sensors in order to limit the dissipation to the region near the shock. Among all discontinuity sensors, the most popular is the use of the gradient ratio. Following the nomenclature in Fig. 2, this is defined as follows [28]:

$$r_f = \frac{\Delta_U \theta}{\Delta_u \theta} = \frac{\theta_C - \theta_U}{\theta_D - \theta_C}, \quad (15)$$

where  $\Delta_U \theta$  is the gradient of  $\theta$  at the upwind face while  $\Delta_u \theta$  correspond with the gradient at the face of interest. Both differences are taken as positive in the flow direction, defined by the sign of the velocity field  $\mathbf{u}_f$ .

This provides with an intuitive description of where discontinuities are and its order of magnitude, at the time that keeps a compact stencil. In addition, it allows to, after proper manipulation by the flux limiter itself, limit the flux in a way which can be interpreted as a diffusion-like term. This is known as ‘‘upwinding’’ as it has the same effect as recovering the 1st order upwind discretization near shocks.

TVD conditions in terms of the gradient ratio were stated by Harten in 1983 [27] for one-dimensional homogeneous grids. These conditions were used by Sweby in 1984 [28] to state 2nd order and TVD conditions for different forms of flux limiters. This idea has been extended, with several degrees of accuracy, to multidimensional and irregular grids [29,30], among others.

#### 3.1. Algebraic formulation

As stated previously, the discretization of equation (14) may benefit from the adoption of an algebraic approach. In this regard, it can be easily extended to the whole computational domain as:

$$\theta_f = (\Pi_{C \rightarrow F} + F(\mathbf{r}_f) Q(\mathbf{u}_f) E_{C \rightarrow F}) \theta_C, \quad (16)$$

where  $\theta_f \in \mathbb{R}^{|F|}$  and  $\theta_C \in \mathbb{R}^{|C|}$  are the vectors holding all the values of  $\theta_f$  and  $\theta_C$ ,  $\Pi_{C \rightarrow F} \in \mathbb{R}^{|C| \times |F|}$  is the cell-to-face interpolation defined in equation (8),  $F(\mathbf{r}_f) \in \mathbb{R}^{|C| \times |F|}$  is the diagonal matrix absorbing the artificial diffusion introduced in equation (14),  $Q(\mathbf{u}_f) \in \mathbb{R}^{|F| \times |F|}$  is the diagonal matrix taking the proper sign of the velocity at the faces and  $E_{C \rightarrow F}$  takes the difference across them as in equation (5).

At this point, we may be tempted to analyze the construction of new flux limiters by means of basic algebra concepts. In particular, to bound its spectrum by means of Gershgorin's theorem or to check its entropy conditions [42], among others. The interested reader is referred to Báez et al. [43], where a similar approach is taken for spatial filters.

While both  $\Pi_{C \rightarrow F}$  and  $E_{C \rightarrow F}$  are readily available from the background stated in section 2, the construction of  $F(\mathbf{r}_f)$  by means of basic algebraic operations solely is addressed next.

Because flux limiter functions  $F(\mathbf{r}_f)$  depend only on the local value of  $\mathbf{r}_f$  and we defer the details on the implementation of the pointwise operations to section 3.2, the problem is turned into the accurate computation of  $\mathbf{r}_f$  at faces. There has been several approaches [31,29] to the construction of  $\mathbf{r}_f$  in terms of a least-squares reconstructed gradient. However, the implementation of such schemes can be cumbersome and may not, eventually, recover the one-dimensional homogeneous solution when a homogeneous structured mesh is used.



The construction of the gradient ratio will proceed first by the separate calculation of both the numerator ( $\Delta_U \theta$ ) and the denominator ( $\Delta_u \theta$ ) of equation (15), then computed as:

$$[\mathbf{r}_f]_i = \frac{[\mathbf{d}_u \theta]_i}{[\mathbf{d}_u \theta]_i}, \quad (17)$$

where  $\mathbf{d}_u \theta \in \mathbb{R}^{|F|}$  is the face-centered vector holding the difference across the face taken in the direction stated by  $\mathbf{u}_f$ , while  $\mathbf{d}_u \theta \in \mathbb{R}^{|F|}$  holds the upstream differences according, again, to  $\mathbf{u}_f$ .

In this approach we propose to employ symmetry-preserving gradients (see [35]) into the calculation of both face-centered and upstream gradients in order to preserve, as much as possible, the mimetic properties of the approach. In addition, we aim at recovering the Cartesian formulation as in [33].

Before any calculation, the sign matrix ( $Q(\mathbf{u}_f)$ ) is constructed by assigning to a  $\mathbb{R}^{|F| \times |F|}$  diagonal matrix +1 for a positive velocity and -1 for a negative one:

$$[\text{diag}(Q(\mathbf{u}_f))]_i = \text{sign}([\mathbf{u}_f]_i). \quad (18)$$

This allows a straightforward calculation of the velocity-oriented gradient at the face as follows:

$$\mathbf{d}_u \theta = Q(\mathbf{u}_f) E_{C \rightarrow F} \theta_c, \quad (19)$$

where  $Q(\mathbf{u}_f)$  is used to provide the right direction in which the difference is taken according to the velocity field.

The construction of the upwind difference  $\mathbf{d}_u \theta$  is more involved. The idea is to construct a partial adjacency matrix which only considers upstream faces, namely the upstream adjacency matrix,  $A_{FF}^U(u)$ , which is responsible to garner upstream information and will be defined further in this paper.

We proceed as follows:  $E_{C \rightarrow F}$  is used to compute the difference across every face according to equation (5). In order to assess the contribution of every neighboring face to the face of interest, differences are vectorized with its corresponding normal using  $N_F^T$  and added all together with  $A_{FF}^U(u)$ . Finally, the resulting value is projected over the normal of the face of interest by means of  $N_F$ . The overall construction of the operator is then:

$$\mathbf{d}_u \theta = N_F (\mathbb{I}_d \otimes A_{FF}^U(u)) N_F^T E_{C \rightarrow F} \theta_c, \quad (20)$$

where, similarly to equation (9), we reuse the  $A_{FF}^U(u)$  operator for all spatial dimensions. Face normal matrices  $N_F$ , defined as in equation (10), are used for both vectorization and projection of the neighboring differences. In this way, orthogonal meshes recover the original one-dimensional formulation, whereas unstructured ones are handled inherently by the incidence matrix.

We are now left with construction of the upstream adjacency matrix,  $A_{FF}^U(u)$ , which may look cumbersome at a first glance. It can be assembled from other simpler matrices:

$$A_{FF}^U(u) = \frac{1}{2} (A_{FF} - Q(\mathbf{u}_f) A_{FF}^D), \quad (21)$$

where  $A_{FF}$  is the face adjacency matrix,  $A_{FF}^D$  is a ‘‘directed adjacency matrix’’, which will be introduced below, and  $Q(\mathbf{u}_f)$  is the already defined velocity sign matrix. The strategy for the construction of  $A_{FF}^U(u)$  is to add the contribution of all neighboring faces irrespective of the flow direction and then use  $Q(\mathbf{u}_f) A_{FF}^D$ , to remove the downwind ones depending on the values of  $Q(\mathbf{u}_f)$ . Note that both  $A_{FF}$  and  $A_{FF}^D$  are constant matrices and that the only matrix that needs to be updated according to  $\mathbf{u}_f$  is  $Q(\mathbf{u}_f)$ .

The construction of  $A_{FF}$  proceeds similarly to equation (12) as follows:

$$A_{FF} = W_{FF} - B_{C \rightarrow F} B_{F \rightarrow C}. \quad (22)$$

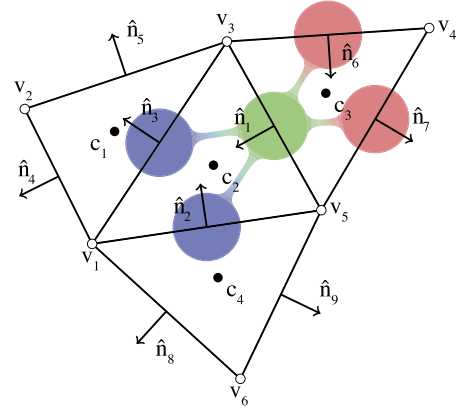


Fig. 3. Upstream (red) and downstream (blue) adjacent faces for face  $f_1$  with respect to the positive component of velocity at face  $f_1$ . The selection of the right ones in  $A_{FF}^U(u)$  will ultimately depend on  $Q(\mathbf{u}_f)$ . (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

As seen in section 2, the adjacency matrix is symmetric and contains non-negative entries only. Following on the example depicted in Fig. 3, the corresponding  $A_{FF}$  reads:

$$A_{FF} = \begin{pmatrix} 0 & +1 & +1 & 0 & 0 & +1 & +1 & 0 & 0 \\ +1 & 0 & +1 & 0 & 0 & 0 & 0 & +1 & +1 \\ +1 & +1 & 0 & +1 & +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & +1 & 0 & 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ +1 & 0 & 0 & 0 & 0 & +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & +1 \\ 0 & +1 & 0 & 0 & 0 & 0 & 0 & +1 & 0 \end{pmatrix}. \quad (23)$$

On the other hand, the construction of  $A_{FF}^D$  allows us to distinguish neighboring faces which lie, according to the face normal, behind or ahead of the face in question. This requires the inclusion of the directed incidence matrix  $E_{C \rightarrow F}$  into the calculation of the adjacency matrix as:

$$A_{FF}^D = E_{C \rightarrow F} B_{F \rightarrow C}, \quad (24)$$

which provides with the following matrix:

$$A_{FF}^D = \begin{pmatrix} 0 & +1 & +1 & 0 & 0 & -1 & -1 & 0 & 0 \\ +1 & 0 & +1 & 0 & 0 & 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & +1 & +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}. \quad (25)$$

Note that  $A_{FF}^D$  has the same pattern as  $A_{FF}$  but entries corresponding to faces located upstream (with respect to the face normal direction) contain  $-1$  whereas those located downstream contain  $+1$ , as shown in Fig. 3.

However, the choice of upstream/downstream faces should depend on the faces local velocity and not on its arbitrary choice of face normal. The product  $Q(\mathbf{u}_f) A_{FF}^D$ , corrects this by inverting the sign of the rows corresponding to the faces whose velocity component is not aligned with the face normal. The result is a correct choice of upstream and downstream faces according to the local face velocity.

Finally, the combination of  $A_{FF}^D$  and  $A_{FF}$  in equation (20) results in:

**Table 1**

Particularizations of the operator  $\circ$  and pointwise function  $f(x_i)$  to represent various kernels using the generalized binary operator described in Equation (28). The `superbeerty` corresponds to the SUPERBEE flux limiter [28].

$\circ$	$f(x_i)$	AI	Resulting Kernel
+	$ax_i$	1/12	<code>axpy</code>
×	$ax_i$	1/12	<code>axty</code>
/	$ax_i$	1/12	<code>axdy</code>
×	$x_i > 0? + 1 : -1$	1/12	<code>signxty</code>
×	$0.5(\max(0, \max(\min(1, 2x_i), \min(x_i, 2))) - 1)$	7/24	<code>superbeerty</code>

$$\mathbf{d}_U \theta = \mathbf{N}_F \left( \mathbb{I}_d \otimes \frac{1}{2} (\mathbf{Q}(\mathbf{u}_f) \mathbf{A}_{FF}^D - \mathbf{A}_{FF}) \right) \mathbf{N}_F^T \mathbf{E}_{C \rightarrow F} \theta_c. \quad (26)$$

While equation (26) succeeds at selecting the proper upstream faces, its direct implementation involves many redundant operations that may result in an unnecessary overhead. For this reason, the computation of  $\mathbf{d}_U \theta$  is rearranged as follows:

$$\mathbf{d}_U \theta = (\mathbf{Q}(\mathbf{u}_f) \mathbf{S}_{C \rightarrow F} + \mathbf{T}_{C \rightarrow F}) \theta_c, \quad (27)$$

where we introduce the new matrices  $\mathbf{S}_{C \rightarrow F} = \frac{1}{2} \mathbf{N}_F (\mathbb{I}_d \otimes \mathbf{A}_{FF}^D) \mathbf{N}_F^T \mathbf{E}_{C \rightarrow F}$  and  $\mathbf{T}_{C \rightarrow F} = \frac{1}{2} \mathbf{N}_F (\mathbb{I}_d \otimes \mathbf{A}_{FF}) \mathbf{N}_F^T \mathbf{E}_{C \rightarrow F}$ , which can be precomputed at the beginning of the simulation.

### 3.2. Algebraic implementation

In previous works of Oyarzun et al. [18] and Álvarez et al. [23], an algebra-based implementation model was proposed for the DNS and LES of incompressible turbulent flows such that the algorithm of the time-integration phase reduces to a set of only three algebraic kernels: `SpMV`, `axpy` and `dot`. However, a close look at Equations (17) and (18), for instance, reveals that this set is insufficient to fulfill the implementation of the flux limiter because it comprises non-linear operations.

Nevertheless, instead of being an inconvenience, this encourages us to demonstrate the high potential of our algebraic strategy again. We propose the generalization of the `axpy` via the introduction of a generalized binary operator (`kbin`) that performs any given pointwise arithmetic calculation such that:

$$y_i \leftarrow y_i \circ f(x_i). \quad (28)$$

This binary operator can easily map to the required kernels by defining  $\circ$  and  $f(x_i)$  as outlined in Table 1. Similarly, the `dot` kernel can be turned into a generalized reduction operator (`kred`),

$$r \leftarrow r \circ f(x_i), \quad (29)$$

which can easily represent any required reduction operation such as the calculation of the norm of a vector or the Courant-Friedrichs-Lewy (CFL) condition. However, the details of the `kred` are out of the scope of this work because it is not required for the implementation of the flux limiter.

From a computational point of view, this kernel generalization does not alter the implementation of the original `axpy`: it still performs simple, pointwise arithmetic operations over the vector elements and provides uniform, aligned and coalescing memory accesses which suits the simple instruction, multiple data (SIMD) and SP paradigms perfectly. Therefore, having already efficient implementations of `axpy` for different architectures, the implementation of `kbin` is straightforward (e.g., consider the use of function pointers, templates, macros, among others).

On the other hand, the arithmetic intensity of this new kernel is not a fixed value anymore, as shown in Table 1. While the AI of the `axpy` was 1/12 FLOP per byte (one product and one addition per three 8-byte values), that of the `kbin` will depend on the specific arithmetic calculations involved in the function  $f(x_i)$ . This allows

us to significantly increase the AI in our calls by means of kernel fusion, reduce the number of intermediate results, and thus reduce the time-to-solution.

The final algorithm for the deployment of a flux limiter in the reconstruction of the variable at faces,  $\theta_f$ , within our algebra-based framework is described in Algorithm 1. Note that because we are

**Algorithm 1** Algorithm for reconstruction of a scalar field at faces,  $\theta_f$ , using the algebraic implementation of a flux limiter.

Require:	$\theta_c, \mathbf{u}_f, \mathbf{E}_{C \rightarrow F}, \mathbf{S}_{C \rightarrow F}, \mathbf{T}_{C \rightarrow F}, \mathbf{N}_{C \rightarrow F}$
Ensure:	$\theta_f$
1:	$\mathbf{d}_U \theta \leftarrow \mathbf{E}_{C \rightarrow F} \theta_c$ <span style="float: right;">▷ SpMV</span>
2:	$\mathbf{d}_U \theta \leftarrow \mathbf{Q}(\mathbf{u}_f) \mathbf{d}_U \theta$ <span style="float: right;">▷ signxty</span>
3:	$\mathbf{r}_f \leftarrow \mathbf{S}_{C \rightarrow F} \theta_c$ <span style="float: right;">▷ SpMV</span>
4:	$\mathbf{r}_f \leftarrow \mathbf{Q}(\mathbf{u}_f) \mathbf{r}_f$ <span style="float: right;">▷ signxty</span>
5:	$\mathbf{r}_f \leftarrow \mathbf{r}_f + \mathbf{T}_{C \rightarrow F} \theta_c$ <span style="float: right;">▷ SpMV</span>
6:	$\mathbf{r}_f \leftarrow \mathbf{r}_f / \mathbf{d}_U \theta$ <span style="float: right;">▷ axdy</span>
7:	$\theta_f \leftarrow \mathbf{F}(\mathbf{r}_f) \mathbf{d}_U \theta$ <span style="float: right;">▷ superbeerty</span>
8:	$\theta_f \leftarrow \theta_f + \mathbf{N}_{C \rightarrow F} \theta_c$ <span style="float: right;">▷ SpMV</span>

actually interested in the evaluation of  $\mathbf{d}_U \theta$  rather than in the construction of the operator itself, matrix-matrix products are avoided and successive matrix-vector products are performed. Similarly, we are not interested in the construction of the diagonal matrices  $\mathbf{Q}(\mathbf{u}_f)$  and  $\mathbf{F}(\mathbf{r}_f)$ . We will rely on the generalized binary operator instead. Therefore, the evaluation of  $\mathbf{Q}(\mathbf{u}_f) \mathbf{y}$  will be done using the `signxty` in Table 1 as follows:

$$[\mathbf{y}]_i \leftarrow [\mathbf{y}]_i \times ([\mathbf{u}_f]_i > 0? + 1 : -1). \quad (30)$$

Finally, a particular flux limiter function must be chosen on the evaluation of  $\mathbf{F}(\mathbf{r}_f) \mathbf{y}$ . Our framework allows to easily switch between different functions. Hereinafter, we will proceed with the `superbee` [28], represented by `superbeerty`:

$$[\mathbf{y}]_i \leftarrow [\mathbf{y}]_i \times 0.5(\max(0, \max(\min(1, 2[\mathbf{r}_f]_i), \min([\mathbf{r}_f]_i, 2))) - 1). \quad (31)$$

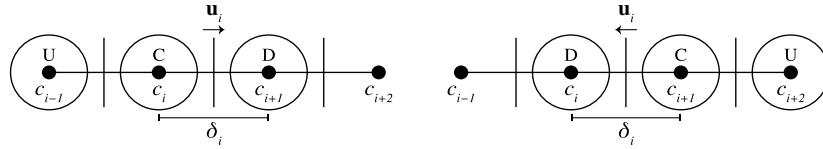
In conclusion, the evaluation of a scalar field at faces,  $\theta_f$ , as in Algorithm 1 can be fitted in our algebra-based framework by combining two types of computing kernels: `SpMV` and `kbin`.

### 3.3. Comparison with stencil-based implementations

Our algebraic implementation for the appropriate reconstruction of the variables at faces,  $\theta_f$ , is now compared with classical, stencil-based approaches. This comparison is conducted firstly from a theoretical point of view, assessing the minimum number of FLOP and memory traffic (in bytes) required in different scenarios. Note that the actual number of memory accesses during kernel execution depends not only on the algorithm but also on hardware and software features. Therefore, regarding the memory traffic, two different values are estimated considering the *full-hit* and *full-miss* caseloads. The former refers to the best scenario with an ideal temporal locality: multiple accesses to a particular data element are so close in time that its value is always reused from cache. Conversely, the latter considers the worst scenario with a null temporal locality so that every repeated access results in cache-miss and requires a memory load from memory. Thus, these two values result in the interval of effective AI of each kernel.

For the sake of clarity, in this comparison we only consider *k-regular*, periodic meshes composed of convex polygons which accomplish the following equality:

$$|F| \approx \frac{k}{2} |C|, \quad (32)$$



**Fig. 4.** Example of stencils in a one-dimensional Cartesian grid according to the classical flux limiter approach. The stencil topology varies according to the sign of the velocity field,  $\vec{v}$ .

where  $|C|$  and  $|F|$  represent the number of cells and faces, respectively, and  $k$  is the degree of the mesh elements (*i.e.*, the number of neighboring cells). The resulting requirements are listed in Table 2, which is described throughout the section.

Let us start from the analysis of the simplest case: the stencil-based calculation of  $\theta_f$  in a one-dimensional Cartesian grid, depicted in Fig. 4 and described in Algorithm 2. Every  $i$ th face is surrounded by two cells,  $c_i$  and  $c_{i+1}$ . For each  $i$ th face in  $|F|$ , the sign of the velocity determines the upstream, centered and downstream values of  $\theta$ , as well as the centered and upstream distances. Following the algorithm,  $5|F|$  FLOP are required for computing the gradient ratio in line 9,  $|F|$  for computing the limiter function in line 10 (this value may vary depending on the limiter function; in this example we have considered the *superbee* limiter [28]), and  $7|F|$  for computing the value at the face in line 11. In the algorithm, three discrete fields are required for the computations: the initial scalar field,  $\theta_c \in \mathbb{R}^{|C|}$ , the velocity field,  $\mathbf{u}_f \in \mathbb{R}^{|F|}$  and the distances between cells,  $\mathbf{d}_f \in \mathbb{R}^{|F|}$ . Besides, the algorithm ensures the calculation of the discrete scalar field at faces,  $\theta_f \in \mathbb{R}^{|F|}$ . Thus, considering double-precision values, and given that the number of faces is equal to the number of cells in the one-dimensional case ( $k = 2 \rightarrow |F| = |C|$ ), the minimum FLOP and bytes required are  $5|F| + 1|F| + 7|F| = 13|C|$  and  $8(|C| + |F| + |F| + |F|) = 32|C|$ , respectively. The total memory traffic in the full-miss caseload would rise to  $48|C|$  because two different values of  $\mathbf{d}_f$  and  $\theta_c$  are accessed for every face. Note that these values are slightly reduced in the particular case of uniform meshes: neither the distances array nor its quotient is required, thus omitting  $2|F|$  FLOP in line 9 and the access to  $\mathbf{d}_f$ . On the other hand, the generalization of Algorithm 2 for three-dimensional Cartesian grids ( $k = 6 \rightarrow |F| = 3|C|$ ) is straightforward and rises the computational requirements to  $39|C|$  FLOP and  $80\text{--}336|C|$  bytes for non-uniform meshes, or  $33|C|$  FLOP and  $56\text{--}216|C|$  bytes in the uniform case.

**Algorithm 2** Stencil-based calculation of  $\theta_f$  in a one-dimensional Cartesian grid.

```

Require:  $\theta_c, \mathbf{u}_f, \mathbf{d}_f$  ▷ 8C + 8F + 8F bytes
Ensure:  $\theta_f$  ▷ 8F bytes
1: for  $i \leftarrow 1$  to  $|F|$  do
2:   if  $\mathbf{u}_f[i] > 0$  then
3:      $\theta_C = \theta_c[i]$ ,  $\theta_U = \theta_c[i - 1]$ ,  $\theta_D = \theta_c[i + 1]$ 
4:      $d_i = \mathbf{d}_f[i]$ ,  $d_U = \mathbf{d}_f[i - 1]$ 
5:   else
6:      $\theta_C = \theta_c[i + 1]$ ,  $\theta_U = \theta_c[i]$ ,  $\theta_D = \theta_c[i + 2]$ 
7:      $d_i = \mathbf{d}_f[i]$ ,  $d_U = \mathbf{d}_f[i + 1]$ 
8:   end if
9:    $r_i \leftarrow (d_i/d_U)(\theta_C - \theta_U)/(\theta_D - \theta_C)$  ▷ 5F flops
10:   $\Psi_i \leftarrow \text{limiter}(r_i)$  ▷ 1F flops
11:   $\theta_f[i] \leftarrow (\theta_C + \theta_D)/2 + (\Psi_i - 1)(\theta_D - \theta_C)/2$  ▷ 7F flops
12: end for

```

A generalization of the stencil calculation for unstructured meshes is outlined in Algorithm 3. In contrast with the structured algorithm, the indices of neighboring nodes are not predictable, so the incidence graphs are required. For each  $i$ th face in  $|F|$ , the sign of the velocity determines the indices of the centered and downstream cells,  $c_C$  and  $c_D$ , according to the cell-to-face incidence graph. Then, for each  $j$ th face incident to  $c_C$  (except  $f_i$ ), its contribution to the upstream gradient, projected over the normal

of  $f_i$ , is accumulated, accounting for  $5(k - 1)|F|$  and  $4(k - 1)|F|$  FLOP in lines 20 and 21, respectively. The calculation of  $\theta_f$  in lines 24 to 26 follows similarly as in Algorithm 2, and adds  $13|F|$  operations. In this case, six discrete fields, one integer list and two incidence graphs are required for the computations. The specific requirements for two  $k$ -regular unstructured meshes,  $k = 2$  (one-dimensional mesh) and  $k = 6$  (three-dimensional hexaedral mesh), are listed in Table 2.

**Algorithm 3** Stencil-based calculation of  $\theta_f$  in a generic unstructured grid based on incidence graphs.

```

Require:  $\theta_c, \mathbf{u}_f, \mathbf{d}_f, \mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z$  ▷ 8C + 40F bytes
         $\mathbf{k}_c, E_{C \rightarrow F}, E_{F \rightarrow C}$  ▷ 4C + 8F + 8F bytes
Ensure:  $\theta_f$  ▷ 8F bytes
1: for  $i \leftarrow 1$  to  $|F|$  do
2:   if  $\mathbf{u}_f[i] > 0$  then
3:      $c_C = E_{C \rightarrow F}[i][0]$ 
4:      $c_D = E_{C \rightarrow F}[i][1]$ 
5:   else
6:      $c_C = E_{C \rightarrow F}[i][1]$ 
7:      $c_D = E_{C \rightarrow F}[i][0]$ 
8:   end if
9:    $\Delta_U = 0$ 
10:  for  $k \leftarrow 1$  to  $\mathbf{k}_c[c_C]$  do
11:     $j = E_{F \rightarrow C}[c_C][k]$ 
12:    if  $j \neq i$  then
13:      if  $E_{C \rightarrow F}[j][0] \neq c_C$  then
14:         $c_U = E_{C \rightarrow F}[j][0]$ 
15:      else
16:         $c_U = E_{C \rightarrow F}[j][1]$ 
17:      end if
18:       $d_j = \mathbf{d}_f[j]$ 
19:       $\theta_U = \theta_c[c_U]$ 
20:       $\mathbf{n}_k \leftarrow \mathbf{n}_x[i]\mathbf{n}_x[j] + \mathbf{n}_y[i]\mathbf{n}_y[j] + \mathbf{n}_z[i]\mathbf{n}_z[j]$  ▷ 5(k-1)F flops
21:       $\Delta_U \leftarrow \Delta_U + \mathbf{n}_k(\theta_C - \theta_U)/d_j$  ▷ 4(k-1)F flops
22:    end if
23:  end for
24:   $r_i \leftarrow \Delta_U / ((\theta_D - \theta_C)/d_i)$  ▷ 3F flops
25:   $\Psi_i \leftarrow \text{superbee}(r_i)$  ▷ 1F flops
26:   $\theta_f[i] \leftarrow (\theta_C + \theta_D)/2 + (\Psi_i - 1)(\theta_D - \theta_C)/2$  ▷ 7F flops
27: end for

```

Finally, in the algebraic implementation outlined in Algorithm 1, it can be observed how it is completely independent of the mesh type and the numerical method: these characteristics only affect the matrices. The number of calls to `SpMV` and `kbin` kernels is readily deduced from the algorithm: 4 times each. Four matrices are required for the computations: the differences at faces,  $E_{C \rightarrow F}$ , with  $2|F|$  non-zero elements, the oriented and un-oriented differences,  $S_{C \rightarrow F}$  and  $T_{C \rightarrow F}$ , with  $2k|F|$  each, and the cell-to-face interpolation,  $\Pi_{C \rightarrow F}$ , with  $2|F|$ . In this example, we consider the use of the `ELLPACK` format [44] in which each non-zero element accounts for 12 bytes (*i.e.*, 8 bytes for the coefficient and 4 bytes for the column index). The specific requirements for two  $k$ -regular meshes are listed in Table 2.

## 4. Numerical study

### 4.1. Three-dimensional deformation problem

Next, the application of this technique is applied to a canonical benchmark. In particular, the deformation (advection) of a sharp

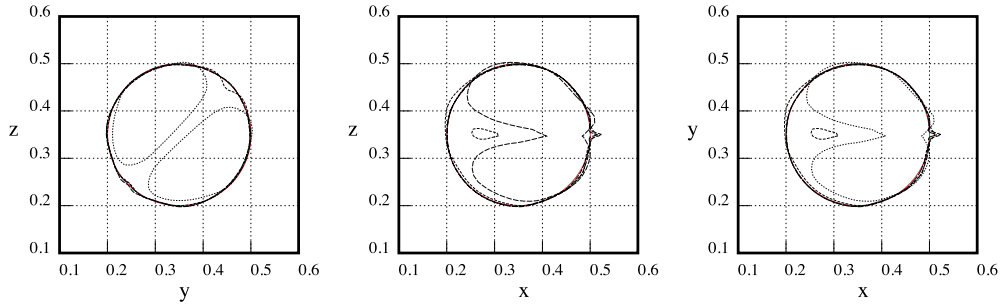


Fig. 5. Contours for  $\theta = 0.5$  in  $x = 0.35$ ,  $y = 0.35$  and  $z = 0.35$  planes after 3.0 time-units for meshes of  $72^3$ ,  $144^3$ ,  $288^3$ ,  $432^3$  and  $576^3$  cells.

Table 2

Minimum number of FLOP and memory traffic (in bytes) required per mesh cell for computing the variable at the faces in different scenarios: stencil- and algebra-based implementations on uniform, non-uniform and unstructured one-dimensional ( $k = 2$ ) and three-dimensional ( $k = 6$ ) meshes.

k	Approach	FLOP	Bytes		AI	
			full-hit	full-miss	full-hit	full-miss
2	Uniform	11	32	48	0.344	0.229
2	Non-uniform	13	40	56	0.325	0.232
2	Unstructured	20	84	132	0.238	0.152
2	Algebraic	28	288	352	0.097	0.080
6	Uniform	33	80	240	0.413	0.138
6	Non-uniform	39	104	360	0.375	0.108
6	Unstructured	168	228	1020	0.737	0.165
6	Algebraic	180	1408	1984	0.128	0.091

profile, which has been tested on three-dimensional hexahedral meshes of  $72^3$ ,  $144^3$ ,  $288^3$ ,  $432^3$  and  $576^3$  cells following Algorithm 4, where  $M \in \mathbb{R}^{|C| \times |F|}$  is the divergence operator [40] and  $U(\mathbf{u}_f) \in \mathbb{R}^{|F| \times |F|}$  is a diagonal matrix containing the velocities at faces. Recall we evaluate the products by diagonal matrices by means of `kbin` calls.

**Algorithm 4** Algorithm for the advection of a scalar field with a 1st order Euler method, using the algebraic implementation of a flux limiter.

**Require:**  $\theta_c^n, \mathbf{u}_f, dt, E_{C \rightarrow F}, S_{C \rightarrow F}, T_{C \rightarrow F}, \Pi_{C \rightarrow F}, M$

**Ensure:**  $\theta_c^{n+1}$

- 1:  $\mathbf{d}_u \theta \leftarrow Q(\mathbf{u}_f) E_{C \rightarrow F} \theta_c^n$
- 2:  $\mathbf{r}_f \leftarrow (Q(\mathbf{u}_f) S_{C \rightarrow F} + T_{C \rightarrow F}) \theta_c^n / \mathbf{d}_u \theta$
- 3:  $\theta_f \leftarrow \Pi_{C \rightarrow F} \theta_c^n + F(\mathbf{r}_f) \mathbf{d}_u \theta$
- 4:  $\theta_c^{n+1} \leftarrow \theta_c^n + dt M U(\mathbf{u}_f) \theta_f$

The sharp profile is initialized in a physical domain of  $[0, 1] \times [0, 1] \times [0, 1]$  as a sphere of radius  $r = 0.15$ , located at  $(0.35, 0.35, 0.35)$  and subject to a divergence-free velocity field:

$$u = 2 \sin^2(\pi x) \sin(2\pi y) \sin(2\pi z) \cos(\pi t/T), \quad (33)$$

$$v = -\sin(2\pi x) \sin^2(\pi y) \sin(2\pi z) \cos(\pi t/T), \quad (34)$$

$$w = -\sin(2\pi x) \sin(2\pi y) \sin^2(\pi z) \cos(\pi t/T), \quad (35)$$

during 3.0 time-units,  $T$  [45].

The results of the profile on meshes of  $72^3$ ,  $144^3$ ,  $288^3$ ,  $432^3$  and  $576^3$  cells are shown in Fig. 5 for the slices in  $x = 0.35$ ,  $y = 0.35$  and  $z = 0.35$  planes.

The three-dimensional temporal evolution of the sphere on a mesh of  $576^3$  cells is shown in Fig. 6. As in [45], the resulting shapes after the deformation are satisfactory, and mass is exactly conserved.

## 4.2. Performance analysis

The benchmark described in Section 4.1 has been deployed on the HPC<sup>2</sup> [23] framework, designed for the efficient implementation of algebraic algorithms on hybrid supercomputers. To demonstrate its portability, the simulations have been run on different supercomputers. Before going into details, we define the *theoretical achievable performance* of a particular kernel,  $\pi_k$ , as:

$$\pi_k = \min(\pi, AI_k \beta),$$

where  $\pi$  is the peak performance of the computing device in double-precision,  $\beta$  is the peak memory bandwidth and  $AI_k$  is the maximum AI of the kernel, taking the full-hit scenario as described in Section 3.3. Then, we define the *performance* and *memory efficiency* as the ratio of measured performance to  $\pi_k$  and measured memory traffic to full-hit, respectively.

The simulations on meshes of  $72^3$ – $432^3$  cells have been executed on up to 64 nodes (3,072 cores) of the CPU-based MareNostrum 4 supercomputer at the Barcelona Supercomputing Center. Its nodes are equipped with two Intel Xeon 8160 CPUs (24 cores, 2.1 GHz, 6 DDR4-2666 memory channels, 128 GB/s memory bandwidth, 33 MB L3 cache), interconnected through the Intel Omni-Path network (12.5 GB/s). The application achieved a sustained performance of up to 1.6 TFLOPS, corresponding to nearly 80% of performance efficiency.

The simulation on a mesh of  $576^3$  cells has been executed on 27 nodes of the Lomonosov-2 hybrid supercomputer at Lomonosov Moscow State University. Its hybrid nodes are equipped with one Intel Xeon E5-2697 v3 CPU (14 cores, 2.6 GHz, 4 DDR4-2133 memory channels, 68 GB/s memory bandwidth, 35 MB L3 cache) and one NVIDIA Tesla K40M GPU (12 GB of GDDR5 memory, 288 GB/s, PCIe 3.0 x16 – 16 GB/s), interconnected via InfiniBand FDR network (7 GB/s). The application achieved a sustained performance of 0.9 TFLOPS, corresponding to nearly 75% of performance efficiency and 98.5% of the heterogeneous efficiency (*i.e.*, the ratio of the heterogeneous performance to the sum of CPU-only and GPU-only performances).

Finally, we complete the performance analysis, and the theoretical comparison in Section 3.3, by comparing the actual performance of the algebraic and stencil approaches. Following Algorithm 3, a stencil kernel has also been implemented in HPC<sup>2</sup>. The tests have been carried out on the CPU-based JFF cluster at the Heat and Mass Transfer Technological Center. Its nodes are equipped with two Intel Xeon Gold 6230 CPUs (20 cores, 2.1 GHz, 6 DDR4-2933 memory channels, 140 GB/s memory bandwidth, 27.5 MB L3 cache). Considering that the DM parallelization is equivalent in both approaches since the data exchanges are the same, only single-node comparisons have been conducted.

The results are shown in a roofline plot in Fig. 7 for both single- and dual-socket executions using a NUMA-aware, SM parallelization on meshes of  $144^3$  (A) and  $288^3$  (B) cells. We have discarded



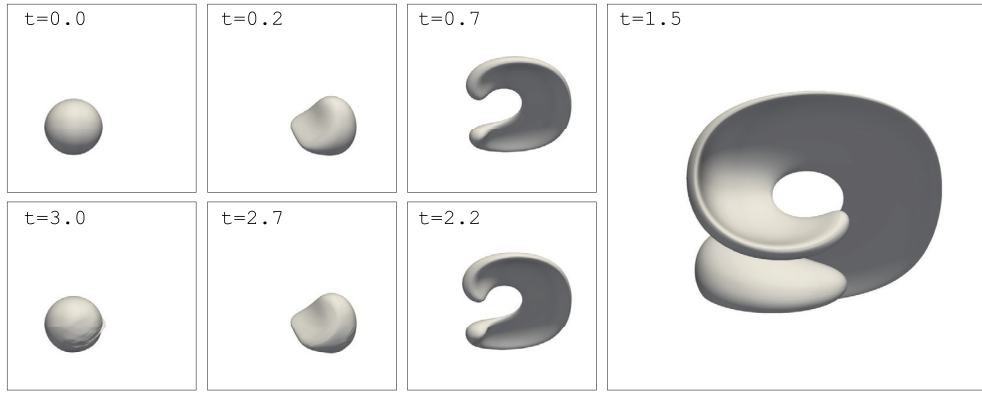


Fig. 6. Time evolution of the  $\theta = 0.5$  contour for  $t = 0, 0.25, 0.75, 1.5, 2.25, 2.75$  and  $3.0$  time-units for mesh of  $576^3$  cells.

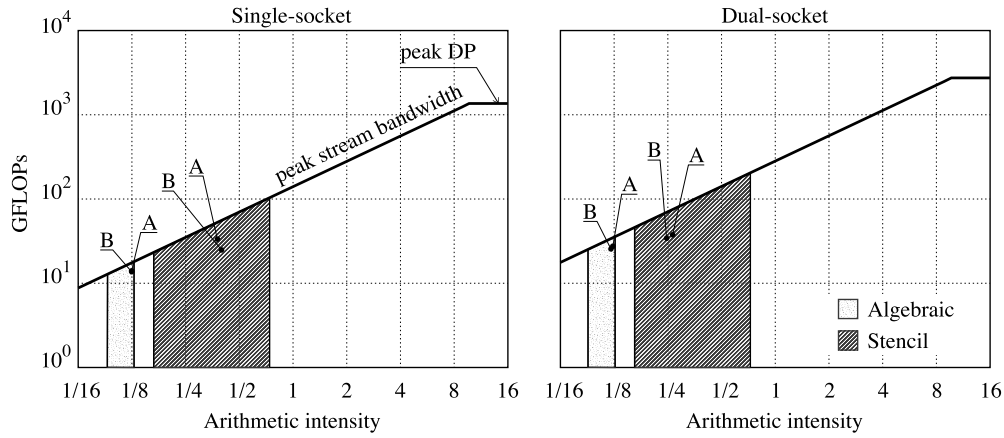


Fig. 7. Roofline models representing the ranges of achievable performance for the stencil- and algebra-based implementations of the flux limiter on three-dimensional unstructured meshes. Results on Intel Xeon Gold 6230 are shown for meshes of  $144^3$  (A) and  $288^3$  (B).

the smallest mesh of  $72^3$  to ensure a memory-bounded behavior and the biggest meshes of  $432^3$  and  $576^3$  because they do not fit in a single node. In the plot, two vertical lines represent the minimum and maximum values of the AI for each kernel as estimated in Section 3.3 and outlined in Table 2. Then, the roofline curve is calculated as follows:

$$\Pi(\pi, \beta) = \min(\pi, A\beta).$$

In this particular test,  $\beta$  and  $\pi$  are 140 GB/s and 1344 GFLOPS per socket, respectively. The real number of memory accesses to main memory have been measured using a profiling tool to calculate the effective AI of each execution.

In single-socket execution, the stencil kernel performs nearly twice faster than the algebraic one (stencil: 33.02 (A) and 24.52 (B) GFLOPS, algebraic: 13.89 (A) and 13.54 (B) GFLOPS), even though its lower performance efficiency (stencil: 32% and 24%, algebraic: 78% and 76%). However, this gap is reduced to approximately  $\times 1.35$  in the dual-socket case (stencil: 37.01 and 33.82 GFLOPS, algebraic: 27.16 and 25.17 GFLOPS). The algebraic kernels feature a regular unit-strided memory access everywhere except the input vector in  $S_{PMV}$ . In contrast, the stencil kernel leads to irregular accesses to  $E_{F \rightarrow C}$ ,  $E_{C \rightarrow F}$ ,  $d_f$ ,  $\theta_c$  and  $\bar{n}$ , resulting in higher cache miss rates and reducing the memory efficiency, especially in dual-socket configurations (stencil: 36% and 33%, algebraic: 96% and 94%). Thus, the actual performance gap is far from the  $\times 5.75$  of the (worst) theoretical scenario.

For further details of the implementation of our framework and a detailed performance and scalability analysis on different types

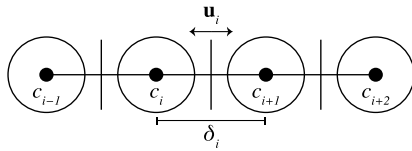
of supercomputing facilities, the reader is referred to Álvarez-Farré et al. [24].

## 5. Discussion

The advantages and disadvantages of the algebra- and stencil-based implementations of the flux limiter are discussed below.

The algebraic formulation of flux limiters proposed in this work allows for fitting the calculation of such high-resolution schemes into algebra-based frameworks. Following this approach, the DNS and LES of turbulent multiphase flows, for instance, is reduced to a minimal set of algebraic subroutines, leading to fully portable and sustainable implementations. The challenges associated with the introduction of new architectures, and the ongoing hybridization of HPC systems make these two properties of the uttermost importance in the development of modern scientific computing codes. Moreover, algebraic kernels are so widespread that optimized libraries are available for virtually all the existing architectures.

From the results listed in Table 2, a significant (theoretical) overhead is revealed in the algebraic implementation. This overhead is mainly induced from Equation (27), where we are enforced to calculate the contribution of all neighboring faces twice as described in Fig. 8, according to the oriented and unoriented matrices, to cancel the downwind values. Note that  $T_{C \rightarrow F}$  and  $S_{C \rightarrow F}$  are the matrices with the largest number of non-zero elements. On the other hand, the algebraic implementation requires more kernel calls which reduces the maximum AI of the algorithm. However, the memory access patterns in algebraic kernels are regular and unit-strided everywhere except the input vector in  $S_{PMV}$ .



**Fig. 8.** Equivalent stencils used in the computation of an algebraic flux limiter. In this case the adjacency matrices involve the operation with all neighboring nodes.

Although the differences in theoretical achievable performance ( $\pi_k$ ) are evident, reaching high performance and memory efficiencies with complex stencil kernels usually requires more complex optimizations. For instance, the authors in [1] study different types of kernels, including  $S_{PMV}$ , and show that simpler kernels require fewer optimizations to reach higher efficiencies, even though their absolute performance is still lower. Also, the authors in [46] study the progressive performance improvement of a CFD solver applying several optimizations such as kernel fusion, cache-blocking, vectorization, and NUMA-aware memory initialization. Indeed, their fully-optimized solver reports absolute performances that are already higher than the  $\pi_k$  of any equivalent algebraic counterpart. However, from their results, the performance gap is far from the (worst) theoretical predictions in Table 2 because their stencil application achieves low performance efficiencies (around 10–40%). In contrast, our measurements in Section 4.2 show that our algebraic kernels report a very stable 80% efficiency. Furthermore, the test case in [46] on a mesh with 2 million grid points appears to be benefiting from cache reuse, especially in the Broadwell device with 56 MB L3 cache, hence the gap on larger grids should be even smaller.

Regarding the parallelization both implementations are very similar. The distributed-memory parallelization remains the same. In both cases, the overlapping schemes for hiding the communication overhead have reported solid results [24,47].

Finally, it is noteworthy that the calculation of  $\theta_f$  represents a marginal part of a simulation, and that the overhead of the algebraic implementation is not significant in other evaluations such as the advection–diffusion of the variables. Indeed, the most time- and memory-consuming part in CFD simulations is the solver of SLAE which, in our implementation, not only does not suffer but benefits: our fields are vectors suitable for the SLAE throughout the entire simulation. Neither copy, transfer, nor manipulation is required for passing our vectors as input parameters to our solvers. Therefore, the drawbacks of the algebraic approach of the flux limiter are diminished in an actual CFD simulation.

## 6. Conclusions and future work

A flux limiter scheme has been formulated from an algebraic perspective resulting in a compact formulation that allows for easy implementation on algebraic frameworks. The resulting implementation provides accurate results and collapses to the traditional approach of Sweby [28] when a homogeneous, Cartesian grid is used.

Graph incidence matrices (both directed and undirected) are exploited to construct appropriate gradient ratios, while the face velocity sign determines the appropriate side to pick the upstream information. After the sign operation, the remaining operations are either linear or local.

This approach presents several advantages, most remarkably reducing the number of computing kernels that need to be ported when moving to new architectures. On the other hand, a theoretical comparison with respect to a classical stencil-based implementation reveals that the latter is cheaper regarding the memory traffic because it can make use of specialized kernels that require less intermediate results, or discriminate some operations with

conditional statements, most remarkably when locating upwind values. However, the performance study shows that the performance gaps are much smaller than the worst theoretical scenario ( $\times 1.35$  instead of  $\times 5.75$ ). Either way, the calculation of  $\theta_f$  represents a marginal part of an actual CFD simulation and, therefore, the drawbacks of the algebraic approach in realistic simulations are diminished.

Finally, the approach developed in this work can be improved to include the effect of the non-homogeneous distance across upstream faces or its surface in calculating the upstream gradient.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The work of N. V. and X. Á. F. has been supported by the Government of Catalonia, FI AGAUR predoctoral grants 2019FI\_B2\_000104 and 2019FI\_B2\_00076.

N. V., X. Á. F., J. C., A. O. and F. X. T. have been funded by the Spanish Research Agency, ANUMESOL project ENE2017-88697-R. J. C. has also been funded by Spanish Research Agency, GALIFLOW project ENE2015-70662-P.

The studies of this work have been carried out using the MareNostrum 4 supercomputer of the Barcelona Supercomputing Center, projects IM-2019-3-0026 and IM-2020-1-0006; the TSUBAME3.0 supercomputer of the Global Scientific Information and Computing Center at Tokyo Institute of Technology; the Lomonosov-2 supercomputer of the shared research facilities of HPC computing resources at Lomonosov Moscow State University; the K-60 hybrid cluster of the collective use center of the Keldysh Institute of Applied Mathematics; the JFF cluster of the Heat and Mass Transfer Technological Center at Technical University of Catalonia. The authors thankfully acknowledge these institutions for the compute time and technical support.

## References

- [1] S. Williams, A. Waterman, D. Patterson, *Commun. ACM* 52 (4) (2009) 65–76, <https://doi.org/10.1145/1498765.1498785>.
- [2] J. Dongarra, M.A. Heroux, P. Luszczek, *Nat. Sci. Rev.* 3 (1) (2016) 30–35, <https://doi.org/10.1093/nsr/nwv084>.
- [3] M. Al Farhan, A. Abdelfattah, S. Tomov, M. Gates, D. Sukkari, A. Haidar, R. Rosenberg, J. Dongarra, *Int. J. High Perform. Comput. Appl.* 34 (6) (2020) 645–658, <https://doi.org/10.1177/1094342020938421>.
- [4] T. Gysi, T. Gresser, T. Hoefler, in: *Proc. 29th ACM Int. Conf. Super-Comput, ACM, New York, NY, USA, 2015*, pp. 177–186.
- [5] B. Krasnopolsky, A. Medvedev, *Adv. Parallel Comput.* 27 (2016) 93102, <https://doi.org/10.3233/978-1-61499-621-7-93>.
- [6] J. Romero, J. Crabill, J.E. Watkins, F.D. Witherden, A. Jameson, *Comput. Phys. Commun.* 250 (2020) 107169, <https://doi.org/10.1016/j.cpc.2020.107169>.
- [7] F. Jiang, K. Matsumura, J. Ohgi, X. Chen, *Comput. Phys. Commun.* 259 (2021) 107661, <https://doi.org/10.1016/j.cpc.2020.107661>.
- [8] S. Watanabe, T. Aoki, *Comput. Phys. Commun.* 264 (2021) 107871, <https://doi.org/10.1016/j.cpc.2021.107871>.
- [9] S. Ha, J. Park, D. You, *Comput. Phys. Commun.* 265 (2021) 107999, <https://doi.org/10.1016/j.cpc.2021.107999>.
- [10] A. Yamanaka, T. Aoki, S. Ogawa, T. Takaki, *J. Cryst. Growth* 318 (1) (2011) 40–45, <https://doi.org/10.1016/j.jcrysgro.2010.10.096>.
- [11] S. Sakane, T. Takaki, M. Ohno, Y. Shibuta, T. Aoki, *Model. Simul. Mater. Sci. Eng.* 27 (5) (2019) 054004, <https://doi.org/10.1088/1361-651X/ab20b9>.
- [12] P. Zaspel, M. Griebel, *Comput. Fluids* 80 (1) (2013) 356–364, <https://doi.org/10.1016/j.compfluid.2012.01.021>.
- [13] X. Zhu, E. Phillips, V. Spandan, J. Donners, G. Ruetsch, J. Romero, R. Ostilla-Mónico, Y. Yang, D. Lohse, R. Verzicco, M. Fatica, R.J. Stevens, *Comput. Phys. Commun.* 229 (2018) 199–210, <https://doi.org/10.1016/j.cpc.2018.03.026>.
- [14] A. Bocharov, N. Evstigneev, V. Petrovskiy, O. Ryabkov, I. Teplyakov, *J. Comput. Phys.* 406 (2020) 109189, <https://doi.org/10.1016/j.jcp.2019.109189>.

- [15] P. Vincent, F. Witherden, B. Vermeire, J.S. Park, A. Iyer, in: *SC16 Int. Conf. High Perform. Comput. Networking, Storage Anal.*, IEEE, Salt Lake City, 2016.
- [16] F. Witherden, A.M. Farrington, P.E. Vincent, *Comput. Phys. Commun.* 185 (11) (2014) 3028–3040, <https://doi.org/10.1016/j.cpc.2014.07.011>.
- [17] T. Shimokawabe, T. Aoki, N. Onodera, *Proc. Comput. Sci.* 80 (2016) 1646–1657, <https://doi.org/10.1016/j.procs.2016.05.499>.
- [18] G. Oyarzun, R. Borrell, A. Gorobets, A. Oliva, *Int. J. Comput. Fluid Dyn.* 31 (9) (2017) 396–411, <https://doi.org/10.1080/10618562.2017.1390084>.
- [19] NVIDIA Corporation, *cuSPARSE: The API reference guide for cuSPARSE, the CUDA sparse matrix library*, Tech. Rep., March 2020.
- [20] J.L. Greathouse, K. Knox, J. Poła, K. Varaganti, M. Daga, in: *Proc. 4th Int. Work. OpenCL*, ACM, New York, NY, USA, 2016.
- [21] G. Oyarzun, R. Borrell, A. Gorobets, A. Oliva, *Comput. Fluids* 92 (2014) 244–252, <https://doi.org/10.1016/j.compfluid.2013.10.035>.
- [22] R. Borrell, J. Chiva, O. Lehmkuhl, G. Oyarzun, I. Rodríguez, A. Oliva, *Int. J. Comput. Fluid Dyn.* 30 (6) (2016) 425–430, <https://doi.org/10.1080/10618562.2016.1221503>.
- [23] X. Álvarez, A. Gorobets, F.X. Trias, R. Borrell, G. Oyarzun, *Comput. Fluids* 173 (2018) 285–292, <https://doi.org/10.1016/j.compfluid.2018.01.034>.
- [24] X. Álvarez-Farré, A. Gorobets, F.X. Trias, *Comput. Fluids* 214 (2021) 104768, <https://doi.org/10.1016/j.compfluid.2020.104768>.
- [25] S. Godunov, *Mat. Sb.* 47 (271) (1959) 271–306.
- [26] C. Hirsch, *Numerical Computation of Internal and External Flows. Vol. 2: Computational Methods for Inviscid and Viscous Flows*, Wiley-Interscience, 1990.
- [27] A. Harten, *J. Comput. Phys.* 49 (3) (1983) 357–393, [https://doi.org/10.1016/0021-9991\(83\)90136-5](https://doi.org/10.1016/0021-9991(83)90136-5).
- [28] P.K. Sweby, *SIAM J. Numer. Anal.* 21 (5) (1984) 995–1011, <https://doi.org/10.1137/0721062>.
- [29] M. Berger, M.J. Aftosmis, S.M. Murman, *Analysis of Slope Limiters on Irregular Grids*, Tech. Rep., 2005, pp. 1–22.
- [30] X. Zeng, *SIAM J. Sci. Comput.* 38 (2) (2016) A789–A813, <https://doi.org/10.1137/140970185>.
- [31] M.S. Darwish, F. Moukalled, *Int. J. Heat Mass Transf.* 46 (4) (2003) 599–611, [https://doi.org/10.1016/S0017-9310\(02\)00330-7](https://doi.org/10.1016/S0017-9310(02)00330-7).
- [32] E. Olsson, G. Kreiss, *J. Comput. Phys.* 210 (1) (2005) 225–246, <https://doi.org/10.1016/j.jcp.2005.04.007>.
- [33] N. Balcázar, L. Jofre, O. Lehmkuhl, J. Castro, J. Rigola, *Int. J. Multiph. Flow* 64 (2014) 55–72, <https://doi.org/10.1016/j.ijmultiphaseflow.2014.04.008>.
- [34] R. Hiemstra, D. Toshniwal, R. Huijsmans, M. Gerritsma, *J. Comput. Phys.* 257 (2014) 1444–1471, <https://doi.org/10.1016/j.jcp.2013.09.027>.
- [35] R.W.C.P. Verstappen, A.E.P. Veldman, *J. Comput. Phys.* 187 (1) (2003) 343–368, [https://doi.org/10.1016/S0021-9991\(03\)00126-8](https://doi.org/10.1016/S0021-9991(03)00126-8).
- [36] N. Valle, F.X. Trias, J. Castro, *J. Comput. Phys.* 400 (2020) 108991, <https://doi.org/10.1016/j.jcp.2019.108991>.
- [37] N. Robidoux, S. Steinberg, *Comput. Methods Appl. Math.* 11 (1) (2011) 23–66, <https://doi.org/10.2478/cmam-2011-0002>.
- [38] K. Lipnikov, G. Manzini, M. Shashkov, *J. Comput. Phys.* 257 (2014) 1163–1227, <https://doi.org/10.1016/j.jcp.2013.07.031>.
- [39] E. Tonti, *J. Comput. Phys.* 257 (2014) 1260–1290, <https://doi.org/10.1016/j.jcp.2013.08.016>.
- [40] F.X. Trias, O. Lehmkuhl, A. Oliva, C.D. Pérez-Segarra, R.W.C.P. Verstappen, *J. Comput. Phys.* 258 (2014) 246–267, <https://doi.org/10.1016/j.jcp.2013.10.031>.
- [41] C. Hirsch, *Numerical Computation of Internal and External Flows. Vol. 1: Fundamentals of Numerical Discretization*, Wiley-Interscience, Brussels, 1988.
- [42] S. Osher, S. Chakravarthy, *High Resolution Schemes and the Entropy Condition*, Tech. Rep., 1984.
- [43] A. Báez Vidal, O. Lehmkuhl, F.X. Trias, C.D. Pérez-Segarra, *J. Comput. Phys.* 326 (4) (2016) 474–498, <https://doi.org/10.1016/j.jcp.2016.09.002>.
- [44] D.R. Kincaid, T.C. Oppe, D.M. Young, *ITPACKV 2D User 's Guide*, Tech. Rep., Center for Numerical Analysis, University of Texas, 1989.
- [45] K. Yang, T. Aoki, *J. Comput. Phys.* 431 (2021) 110113, <https://doi.org/10.1016/j.jcp.2021.110113>.
- [46] B. Mostafazadeh, F. Marti, F. Liu, A. Chandramowlishwaran, in: *Proc. - 2018 IEEE 32nd Int. Parallel Distrib. Process. Symp. IPDPS 2018*, 2018, pp. 753–762.
- [47] A. Gorobets, S. Soukov, P. Bogdanov, *Comput. Fluids* 173 (2018) 171–177, <https://doi.org/10.1016/j.compfluid.2018.03.011>.