INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com/journal/nmf). DOI: 10.1002/fld.3688

A formally second-order cell centred scheme for convection–diffusion equations on general grids

L. Piar[1], F. Babik[1], R. Herbin[2] and J.-C. Latché[1,*,†]

[1]*Institut de Radioprotection et de Sûreté Nucléaire (IRSN), Cadarache, France*
[2]*LATP, Université de Provence, Marseille, France*

SUMMARY

We propose, in this paper, a finite volume scheme to compute the solution of the convection–diffusion equation on unstructured and possibly non-conforming grids. The diffusive fluxes are approximated using the recently published SUSHI scheme in its cell centred version, that reaches a second-order spatial convergence rate for the Laplace equation on any unstructured two-dimensional/three-dimensional grids. As in the MUSCL method, the numerical convective fluxes are built with a prediction-limitation process, which ensures that the discrete maximum principle is satisfied for pure convection problems. The limitation does not involve any geometrical reconstruction, thus allowing the use of completely general grids, in any space dimension. Copyright © 2012 John Wiley & Sons, Ltd.

Received 11 February 2011; Revised 30 March 2012; Accepted 16 April 2012

KEY WORDS: finite volumes; MUSCL method; convection dominant regime; discrete maximum principle; convergence analysis

1. INTRODUCTION

In this paper, we address the following convection–diffusion problem:

$$\partial_t \bar{u} + \operatorname{div}(\bar{u}\boldsymbol{v}) - \operatorname{div}(\kappa \nabla \bar{u}) = f \qquad \text{on } \Omega \times (0,T), \tag{1a}$$

$$\bar{u}(\boldsymbol{x},t) = \bar{u}_{\mathrm{D}}(\boldsymbol{x},t) \qquad \text{on } \partial\Omega_{\mathrm{D}} \times (0,T), \tag{1b}$$

$$-\kappa \nabla \bar{u}(\boldsymbol{x},t) \cdot \boldsymbol{n} = g_{\mathrm{N}}(\boldsymbol{x},t) \qquad \text{on } \partial\Omega_{\mathrm{N}} \times (0,T), \tag{1c}$$

$$\bar{u}(\boldsymbol{x},0) = \bar{u}_0(\boldsymbol{x}) \qquad \text{in } \Omega, \tag{1d}$$

where $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, is an open connected and bounded domain supposed to be polygonal ($d = 2$) or polyhedral ($d = 3$), $\bar{u}$ is the unknown, $\boldsymbol{x}$ and t stand for space coordinates and time, $\boldsymbol{v}$ is a divergence-free velocity field, κ is a non-negative constant real number and $f \in \mathrm{L}^2(\Omega \times (0,T))$ is a given source term. The boundary $\partial\Omega$ of Ω is split into $\partial\Omega_{\mathrm{D}}$ and $\partial\Omega_{\mathrm{N}}$ (the so-called Dirichlet and Neumann boundaries), which form a partition of $\partial\Omega$, and satisfy the following: (i) that the measure of $\partial\Omega_{\mathrm{D}}$ is positive; and (ii) that $\boldsymbol{v} \cdot \boldsymbol{n} \geqslant 0$ on $\partial\Omega_{\mathrm{N}}$, where $\boldsymbol{n}$ is the unit normal vector to $\partial\Omega$ outward Ω. If $\kappa = 0$, we suppose in addition that $\boldsymbol{v} \cdot \boldsymbol{n} \leqslant 0$ on $\partial\Omega_{\mathrm{D}}$, that is, as usual for the transport equation, that $\partial\Omega_{\mathrm{D}}$ and $\partial\Omega_{\mathrm{N}}$ are respectively the inflow and outflow parts of the boundary. The functions $\bar{u}_0$, $\bar{u}_{\mathrm{D}}$ and g_{N} denote the initial value, the Dirichlet boundary value and the diffusion flux prescribed at the Neumann boundary, respectively.

*Correspondence to: J.-C. Latché, Institut de Radioprotection et de Sûreté Nucléaire (IRSN), Cadarache, France.
†E-mail: jean-claude.latche@irsn.fr

In this paper, we propose a cell-centred scheme for the solution of (1), able to cope with almost arbitrary meshes, including non conforming meshes.

The diffusion term is discretized by a cell-centred scheme, which was first presented in [1] and tested in [2] for oil engineering simulation problems. Its convergence analysis was performed in [3] in the framework of its more general version, Scheme Using Stabilization and Hybrid Interfaces (SUSHI) for the discretization of anisotropic and heterogeneous diffusion problems on general non-conforming grids. This scheme is implemented here with a minor variant for the approximation of Neumann boundary conditions, which completely eliminates the unknowns at the faces of the mesh (including boundary faces), thus restricting the set of the unknowns to the cell values only.

For the convection term, we use a finite volume approach, based (as in MUSCL-type schemes) on a two-step algorithm: we first compute a tentative approximation of the unknown field at the face by an affine reconstruction, then modify it (by a so-called limitation procedure) to ensure the L^∞-stability of the scheme. The development of this type of schemes has been the subject of a huge amount of literature; we refer to [4–6] for seminal works for one-dimensional (1D) problems, [7] for a review of the adaptation of these ideas in multi-dimensional spaces, and [8–11] for recent works. In most approaches, the limitation procedure is presented as a limitation of the slope defined by the cell and face values, on the basis of its comparison with other slopes defined by the values taken by the unknown in the neighbourhood. Then, under geometric assumptions for the mesh, this limitation may be shown to imply some conditions (let us call them *stability conditions*) for the approximation at the face, which ensure, for pure convection problems, a local maximum principle [12, 13]. Our strategy here (see also [14] for an ongoing related work) is based on the following remark: for a linear convection term, these *stability conditions* may be exploited to define an admissible interval for the value at the face. This suggests a crude limitation process, which does not use any slope computation and simply consists in performing a (1D) projection of the tentative affine reconstructed face value on this interval. In addition, *stability conditions* are purely algebraic (in the sense that they do not require any geometric computation), and thus work with arbitrary meshes.

This paper is organized as follows. We first introduce some definitions and notations for the mesh (Section 2). We then describe the scheme (Section 3), the discretization of the diffusion and convection terms being detailed in Section 3.1 and Section 3.2, respectively. We conclude the paper with some numerical tests in Section 4.

2. THE MESH

A finite volume discretization of Ω is defined by a triplet $(\mathcal{M}, \mathcal{E}, \mathcal{P})$, where

- $\mathcal{M}$ is a set of non–empty convex open disjoint subsets K of Ω (the control volumes), such that $\bar{\Omega} = \bigcup_{K\in\mathcal{M}} \bar{K}$.
- $\mathcal{E}$ is the set of edges (in two space dimension (2D)) or faces (in three space dimension (3D)), denoted by σ. We denote by $\mathcal{E}(K) \subset \mathcal{E}$ the set of faces of $K \in \mathcal{M}$, by $\mathcal{E}_{\text{ext}}$ and $\mathcal{E}_{\text{int}}$ the set of boundary and interior faces, respectively. The set $\mathcal{E}_{\text{ext}}$ is split in two disjoint subsets, $\mathcal{E}_{\text{ext}} = \mathcal{E}_{\text{D}} \cup \mathcal{E}_{\text{N}}$, where $\mathcal{E}_{\text{D}}$ (respectively $\mathcal{E}_{\text{N}}$) stands for the set of boundary faces included in $\partial\Omega_{\text{D}}$ (respectively $\partial\Omega_{\text{N}}$). Each internal edge $\sigma \in \mathcal{E}_{\text{int}}$ is supposed to have exactly two neighbouring cells, say K, $L \in \mathcal{M}$, and $\bar{K} \cap \bar{L} = \bar{\sigma}$, which we also write $\sigma = K|L$.
- $\mathcal{P} = (\boldsymbol{x}_K)_{K\in\mathcal{M}}$ is a set of points of Ω such that, $\forall K \in \mathcal{M}$, $\boldsymbol{x}_K \in K$.

We will need hereafter the following definitions. The normal vector to a face σ of K outward, K is denoted by $\boldsymbol{n}_{K,\sigma}$, and $\boldsymbol{x}_\sigma$ stands for the mass centre of the face σ. For any $K \in \mathcal{M}$ and any face $\sigma \in \mathcal{E}(K)$, we define the volume $D_{K,\sigma}$ as the cone with basis σ and vertex $\boldsymbol{x}_K$. For $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}$, we denote by $|K|$ the measure of K and by $|\sigma|$ the $(d-1)$-measure of the face σ (Figure 1).

Finally, for the time discretization, we use a constant time step, denoted by δt, and we define $t^n = n\,\delta t$, for $0 \leqslant n \leqslant N = T/\delta t$.

DOI: 10.1002/fld

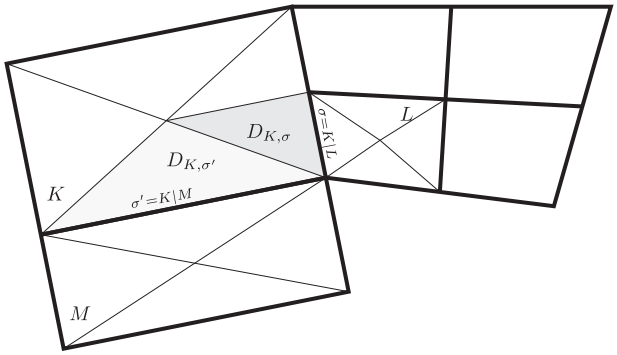


Figure 1. Notations for control volumes and diamond cells.

3. THE SCHEME

The discretization of (1) is performed by a first-order Euler scheme. It combines an explicit discretization of the convection operator and an implicit discretization of the diffusion term. Denoting by $u^n = (u_K^n)_{K\in\mathcal{M}}$ the discrete unknowns at time t^n, $1 \leqslant n \leqslant N$, the scheme reads:

for $0 \leqslant n \leqslant N-1$, $\forall K \in \mathcal{M}$,

$$\frac{|K|}{\delta t}\left(u_K^{n+1} - u_K^n\right) + \sum_{\sigma\in\mathcal{E}(K)} F_{K,\sigma} u_\sigma^n - \kappa\,|K|\,\left(\Delta_{\mathcal{M}} u^{n+1}\right)_K = \frac{1}{\delta t}\int_{t^n}^{t^{n+1}} \int_K f(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \quad (2)$$

with, for $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}(K)$

$$u_K^0 = \frac{1}{|K|}\int_K \bar{u}_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \qquad F_{K,\sigma} = \frac{1}{\delta t}\int_{t^n}^{t^{n+1}} \int_\sigma \boldsymbol{v}(\boldsymbol{x},t)\cdot\boldsymbol{n}_{K,\sigma}\,\mathrm{d}\gamma(\boldsymbol{x})\,\mathrm{d}t.$$

Note that, because $\boldsymbol{v}$ is divergence free, we obtain

$$\forall K \in \mathcal{M}, \quad \sum_{\sigma\in\mathcal{E}(K)} F_{K,\sigma} = 0. \quad (3)$$

The next two subsections are devoted to the description of the discrete diffusion operator ($\Delta_{\mathcal{M}} u^{n+1}$ in (2)) and the discrete convection term, which consists in the choice of the value of u_σ^n in (2).

3.1. Discretization of the diffusion operator

The idea for the discretization of the diffusion operator is related to that of the Galerkin methods; dropping for readability reasons the exponent $n+1$ referring to time, it consists in exploiting an expression of the following form:

$$(-\Delta_{\mathcal{M}} u)_K = \frac{1}{|K|}\left[\int_\Omega \nabla_{\mathcal{M}} u \cdot \nabla_{\mathcal{M}} \mathbf{1}^K\,\mathrm{d}\boldsymbol{x} - \int_{\partial\Omega_{\mathrm{N}}} g_{\mathrm{N}}\,\mathbf{1}^K\,\mathrm{d}\boldsymbol{x}\right], \quad (4)$$

where g_{N} stands for the flux at the Neumann boundary, $\mathbf{1}^K$ is a characteristic function associated to the cell K and $\nabla_{\mathcal{M}}$ denotes an ad hoc discrete gradient operator, which we define in this section.

We start by choosing, for any internal $\sigma \in \mathcal{E}_{\mathrm{int}}$ and any face of the Neumann boundary $\sigma \in \mathcal{E}_{\mathrm{N}}$, some real coefficients $(\beta_\sigma^L)_{L\in\mathcal{M}}$ such that the mass centre $\boldsymbol{x}_\sigma$ of σ is expressed by

$$\boldsymbol{x}_\sigma = \sum_{L\in\mathcal{M}} \beta_\sigma^L \boldsymbol{x}_L, \qquad \sum_{L\in\mathcal{M}} \beta_\sigma^L = 1. \quad (5)$$

Of course, the cells involved in this (non-necessarily convex) interpolation must be chosen as close as possible to the face σ, to reduce the stencil of the scheme. The cells that are adjacent to σ are

DOI: 10.1002/fld

always used; so, for an internal face $K|L$, if $\boldsymbol{x}_K$, $\boldsymbol{x}_\sigma$ and $\boldsymbol{x}_L$ are aligned, only two coefficients in the set $(\beta_\sigma^M)_{M\in\mathcal{M}}$ (namely β_σ^K and β_σ^L) differ from zero. In any case, it is always possible to restrict the number of nonzero coefficients to three in two-space dimensions and to four in three-space dimensions. In practice, we also try to avoid large variations of their values. The algorithm implemented to evaluate these quantities in the tests presented here is given in Remark 7.

We then introduce a second-order interpolation operator of a discrete function at the points $(\boldsymbol{x}_\sigma)_{\sigma\in\mathcal{E}}$. In usual finite element formulations, the Dirichlet boundary conditions are incorporated in the definition of the discrete space; here, although the unknowns are piecewise constant, the analogous effect is obtained by taking Dirichlet boundary conditions into account in the definition of the interpolation operator. This latter thus acts on a larger set than that of the cell values; we call this set a discrete family, defined as the union of cell values and values at the centres of mass of the Dirichlet faces: $w_\mathcal{M} = ((w_K)_{K\in\mathcal{M}}, (w_{\mathrm{D},\sigma})_{\sigma\in\mathcal{E}_\mathrm{D}})$. Then the interpolate of the family $w_\mathcal{M}$ is defined by the data of its values at every face of the mesh $(\tilde{w}_\sigma)_{\sigma\in\mathcal{E}}$:

$$\left|\begin{array}{ll} \forall\sigma\in\mathcal{E}_\mathrm{int}\cup\mathcal{E}_\mathrm{N}, & \tilde{w}_\sigma = \displaystyle\sum_{K\in\mathcal{M}} \beta_\sigma^K w_K, \\ \forall\sigma\in\mathcal{E}_\mathrm{D}, & \tilde{w}_\sigma = w_{\mathrm{D},\sigma}. \end{array}\right. \tag{6}$$

We may then introduce, for any discrete family $w_\mathcal{M}$, a first gradient $\overline{\nabla} w_\mathcal{M}$, defined by its constant value $(\overline{\nabla} w_\mathcal{M})_K$ on each cell K:

$$\forall K\in\mathcal{M},\quad (\overline{\nabla} w_\mathcal{M})_K = \frac{1}{|K|}\sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,(\tilde{w}_\sigma - w_K)\,\boldsymbol{n}_{K,\sigma}. \tag{7}$$

The discrete gradient thus defined is consistent, thanks to the following geometrical identity:

$$\forall K\in\mathcal{M},\quad \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,\boldsymbol{n}_{K,\sigma}\,(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)^t = |K|\,\boldsymbol{I}, \tag{8}$$

where $\boldsymbol{I}$ denotes the identity matrix. Indeed, let ψ be an affine function: $\boldsymbol{x}\mapsto \boldsymbol{a}\cdot\boldsymbol{x}+\boldsymbol{b}$, with $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^d$ (so that $\nabla\psi = \boldsymbol{a}$). Let $\psi_\mathcal{M}$ be the discrete family defined by $\psi_K = \psi(\boldsymbol{x}_K)$ for $K\in\mathcal{M}$ and $\psi_{\mathrm{D},\sigma} = \psi(\boldsymbol{x}_\sigma)$ for any $\sigma\in\mathcal{E}_\mathrm{D}$, and let $\overline{\nabla}\psi_\mathcal{M}$ denote its discrete gradient defined by (7). Because ψ is affine, by definition of a second-order interpolation, we have $\psi_\sigma - \psi_K = \psi(\boldsymbol{x}_\sigma) - \psi(\boldsymbol{x}_K) = \boldsymbol{a}\cdot(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)$. Therefore,

$$(\overline{\nabla}\psi_\mathcal{M})_K = \frac{1}{|K|}\sum_{\sigma\in\mathcal{E}_K} |\sigma|\;(\boldsymbol{a}\cdot(\boldsymbol{x}_\sigma - \boldsymbol{x}_K))\;\boldsymbol{n}_{K,\sigma} = \boldsymbol{a},$$

thanks to (8), and the discrete gradient of the interpolate of an affine function is equal to its exact gradient.

Unfortunately, this is not sufficient to ensure the convergence of the scheme. We also need a weak convergence property [3] and a stability property, which is not satisfied by the previously mentioned gradient: indeed, as noted in [3] and illustrated in [15] in the case of Cartesian grids, this discrete gradient may vanish for nonzero functions. We are thus lead to introduce the following stabilization term, defined for any discrete family $w_\mathcal{M}$ by

$$\forall K\in\mathcal{M},\ \forall\sigma\in\mathcal{E}(K),\quad (Rw_\mathcal{M})_{K,\sigma} = \frac{\sqrt{d}}{d(\boldsymbol{x}_K,\sigma)}\left[\tilde{w}_\sigma - w_K - (\overline{\nabla} w)_K\cdot(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\right], \tag{9}$$

where $d(\boldsymbol{x}_K,\sigma)$ stands for the distance of $\boldsymbol{x}_K$ to σ. Note that $(Rw_\mathcal{M})_{K,\sigma}$ vanishes if $w_\mathcal{M}$ is the interpolate of an affine function; the quantity $(Rw_\mathcal{M})_{K,\sigma}$ may thus be seen as a consistency error on the half-diamond cell $D_{K,\sigma}$.

We then define the discrete gradient of a family $w_\mathcal{M}$ as the piecewise constant function on the half-diamond cells $D_{K,\sigma}$ defined by

$$(\nabla w_\mathcal{M})_{K,\sigma} = (\overline{\nabla} w_\mathcal{M})_K + (Rw_\mathcal{M})_{K,\sigma}\,\boldsymbol{n}_{K,\sigma},\quad \text{on } D_{K,\sigma}. \tag{10}$$

DOI: 10.1002/fld

We can now define the diffusion term $(-\Delta_{\mathcal{M}} u)_K$, for $K \in \mathcal{M}$. To the unknown u, we associate the discrete family $u_{\mathcal{M}} = ((u_K)_{K\in\mathcal{M}}, (u_{D,\sigma})_{\sigma\in\mathcal{E}_D})$, where $u_{D,\sigma}$ stands for the mean value over the face σ of the Dirichlet condition $\bar{\boldsymbol{u}}_D$. Then we define $\mathbf{1}^K$ as the discrete family defined by $(\mathbf{1}^K)_K = 1$, $(\mathbf{1}^K)_L = 0$ for any cell $L \neq K$, and $(\mathbf{1}^K)_{D,\sigma} = 0,\ \forall \sigma \in \mathcal{E}_D$. The value of $(-\Delta_{\mathcal{M}} u)_K$ is then given by (4), with the definition (10) of the discrete gradient, thus, specifying the domains of integration:

$$(-\Delta_{\mathcal{M}} u)_K = \frac{1}{|K|}\left[\sum_{L\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}(L)} |D_{L,\sigma}|\,(\nabla u_{\mathcal{M}})_{L,\sigma}\cdot\left(\nabla \mathbf{1}^K\right)_{L,\sigma} - \sum_{\sigma\in\mathcal{E}_N}\left(\mathbf{1}^K\right)_\sigma \int_\sigma g_N \,\mathrm{d}\gamma(\boldsymbol{x})\right],$$

where $\left(\mathbf{1}^K\right)_\sigma$ is given by (6).

Remark 1 (Consistency with the two-point scheme and maximum principle)
If one only wants to ensure the stability of the scheme, the quantity $(Ru_{\mathcal{M}})_{K,\sigma}$ is defined up to a multiplicative constant; indeed the specific coefficient $\sqrt{d}$ in (9) is chosen so as to recover the usual finite volume two-point diffusion flux for (2D) acute angle triangular meshes and for rectangular grids ($d = 2$ or $d = 3$), provided that the choice for the points $(\boldsymbol{x}_K)_{K\in\mathcal{M}}$ is the usual one, namely the circumcentre of the triangle K in the first case and the mass centre of K in the second one [3, Lemma 2.1].

In this case, the proposed discretization of the diffusion term thus satisfies a discrete maximum principle, which does not hold in the general case, as stated in Theorem 3.2.

Remark 2 (Extension to variable diffusion coefficients)
The extension of the scheme to variable diffusion coefficient may be carried out by defining a diffusion coefficient κ_K, for $K \in \mathcal{M}$, and changing (4) to:

$$((-\mathrm{div}\kappa\nabla)_{\mathcal{M}} u)_K = \frac{1}{|K|}\left[\int_\Omega \kappa\ \nabla_{\mathcal{M}} u\cdot\nabla_{\mathcal{M}}\mathbf{1}^K \,\mathrm{d}\boldsymbol{x} - \int_{\partial\Omega_N} g_N\,\mathbf{1}^K\,\mathrm{d}\gamma(\boldsymbol{x})\right].$$

The result of Remark 1 then still holds, with an expression of the (two-point) flux involving a diffusion at the face which is identical to the classical harmonic average under geometrical conditions on the mesh (see [3, Lemma 2.1] for the expression of the diffusion coefficient).

Remark 3 (The interpolate of the functions 1^K in 1D)
For the sake of clarity, we illustrate the construction of the interface values. Let us suppose that we work on the 1D domain $\Omega = (0, 1)$, with a constant space step. In addition, we suppose that the solution is prescribed at $\boldsymbol{x} = 0$ and obeys a Neumann boundary condition at $\boldsymbol{x} = 1$, so that we have to provide a way to calculate an interpolated value of discrete families at any internal interface and at the interface located at $\boldsymbol{x} = 1$. To this purpose, a reasonable choice seems to be the following:

- at an internal face, the interpolated value is defined as the average of the values taken at the two neighbouring cells;
- at the interface σ located at $\boldsymbol{x} = 1$, we set

$$\tilde{u}_\sigma = \frac{3}{2}u_{K_N} - \frac{1}{2}u_{K_{N-1}},$$

  where N stands for the number of cells, and the cells are indexed from $\boldsymbol{x} = 0$ to $\boldsymbol{x} = 1$.

The interpolated values obtained with this choice for the characteristic function of the cells of the mesh are given in Figure 2.

3.2. *Discretization of the convection operator*

The strategy used to design the convection scheme relies on the following remark: it is possible to state some conditions for the values $(u_\sigma^n)_{\sigma\in\mathcal{E}}$ approximating the unknown at the faces in the convection operator, which are sufficient to ensure that the scheme satisfies a discrete maximum principle

DOI: 10.1002/fld

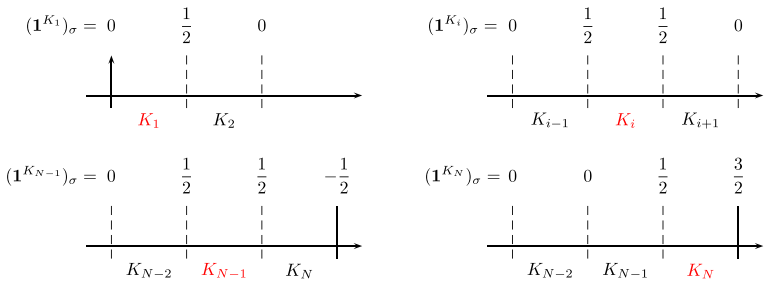


Figure 2. Interpolated values for the characteristic functions of the cells in 1D.

(in a sense given by Lemma 3.1 and Theorem 3.2); moreover, these conditions provide an admissible interval for the $(u_\sigma^n)_{\sigma\in\mathcal{E}}$. Then, a discretization naturally follows: first, compute a tentative value for $(u_\sigma^n)_{\sigma\in\mathcal{E}}$ by an affine interpolation, and then 'limit the flux' (according to the terminology of the MUSCL family of schemes) by projecting this value on the admissible interval.

The exposition follows this line: we first state the conditions to be satisfied by the face values (Section 3.2.1), then show how they may be exploited to obtain a limitation procedure (Section 3.2.2).

3.2.1. Conditions for the satisfaction of a maximum principle. The usual first-order scheme for the convection operator is the upstream scheme, which consists in choosing u_σ^n in (2) for an internal face $\sigma = K|L$ as follows:

$$u_\sigma^n = u_K \text{ if } F_{K,\sigma} \geqslant 0, \quad u_\sigma^n = u_L \text{ otherwise.} \tag{11}$$

The upstream choice is well-known to ensure the maximum principle. Let us briefly review the ingredients which yield this property. For $K \in \mathcal{M}$, let $\tilde{u}_K^{n+1}$ stand for the value updated with the convection term:

$$\tilde{u}_K^{n+1} = u_K^n - \frac{\delta t}{|K|} \sum_{\sigma\in\mathcal{E}(K)} F_{K,\sigma}\, u_\sigma^n. \tag{12}$$

Using the discrete divergence-free constraint (3), with u_σ^n defined by (11), we obtain the following:

$$\tilde{u}_K^{n+1} = \left[1 - \frac{\delta t}{|K|} \sum_{\sigma\in\mathcal{E}(K)} F_{K,\sigma}^+\right] u_K^n + \frac{\delta t}{|K|} \sum_{\sigma\in\mathcal{E}(K),\sigma=K|L} F_{K,\sigma}^-\, u_L^n, \tag{13}$$

where, for any real number a, $a^+ = \max(a,0)$ and $a^- = -\min(a,0)$ (so that $a = a^+ - a^-$). We denote by cfl the following number:

$$\text{cfl} = \max_{K\in\mathcal{M}} \left\{ \frac{\delta t}{|K|} \sum_{\sigma\in\mathcal{E}(K)} |F_{K,\sigma}| \right\}. \tag{14}$$

Then, under the so-called Courant–Friedrichs–Lewy (CFL) condition cfl $\leqslant 1$, Equation (13) yields that $\tilde{u}_K^{n+1}$ is a convex combination of the values taken by u^n in the neighbouring cells of K.

The same principle holds for MUSCL-type schemes, which attempt to go higher order while remaining stable: a piecewise linear reconstruction is performed to evaluate the quantities u_σ^n, and the resulting slopes are limited in order for the expression (12) to remain a convex combination of the values taken by u^n in the neighbouring cells of K. For a linear advection term such as addressed

DOI: 10.1002/fld

here, this latter property is shown to be ensured by the following conditions. First, for internal faces, we suppose

$$\forall K \in \mathcal{M},\ \forall \sigma \in \mathcal{E}(K) \cap \mathcal{E}_{\rm int},\ \text{there exists}\, \alpha_\sigma^K \in [0,1] \text{ and a cell } M_\sigma^K \in \mathcal{M} \text{ such that}$$
$$u_\sigma^n - u_K^n = \begin{cases} \alpha_\sigma^K \left(u_K - u_{M_\sigma^K}\right) & \text{if } F_{K,\sigma} \geqslant 0 \\ \alpha_\sigma^K \left(u_{M_\sigma^K} - u_K\right) & \text{otherwise.} \end{cases} \tag{15}$$

Then, we denote by $\mathcal{E}_{\rm D}^-$ (resp. $\mathcal{E}_{\rm D}^+$) the faces σ of $\mathcal{E}_{\rm D}$ where the flow is entering (resp. leaving) the domain, that is, $F_{K,\sigma} \leqslant 0$ (resp. $F_{K,\sigma} \geqslant 0$). For the faces included in $\mathcal{E}_{\rm D}^+$ and $\mathcal{E}_{\rm N}$ (where, by assumption $F_{K,\sigma} \geqslant 0$), we suppose that the first part of (15) holds. For faces of $\mathcal{E}_{\rm D}^-$, we suppose that u_σ^n is given by the boundary conditions, which we denote by $u_\sigma^n = u_{{\rm D},\sigma}^n$.

The obtained stability property is stated later, and its proof is recalled for the sake of completeness.

Lemma 3.1
Let us suppose that cfl $\leqslant 1$. Let $K \in \mathcal{M}$. We denote by $\mathcal{N}_{\rm m}(K)$ the set of cells M_σ^K, $\sigma \in \mathcal{E}(K)$, which are such that (15) is satisfied. Then, $\forall K \in \mathcal{M}$, the value $\tilde{u}_K^{n+1}$ given by (12) is a convex combination of $\{u_K^n,\ (u_M^n)_{M \in \mathcal{N}_{\rm m}(K)},\ (u_{{\rm D},\sigma}^n)_{\sigma \in \mathcal{E}_{\rm D}^- \cap \mathcal{E}(K)}\}$.

Proof
Let $K \in \mathcal{M}$. By definition, we obtain

$$\frac{|K|}{\delta t}\tilde{u}_K^{n+1} = \frac{|K|}{\delta t}u_K^n - \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}\, u_\sigma^n,$$

and thus, invoking the discrete divergence-free constraint (3):

$$\frac{|K|}{\delta t}\tilde{u}_K^{n+1} = \frac{|K|}{\delta t}u_K^n - \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^+ \left(u_\sigma^n - u_K^n\right) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^- \left(u_\sigma^n - u_K^n\right).$$

Let us first consider the internal faces where the flow is entering K, that is, $F_{K,\sigma}^- \geqslant 0$, $F_{K,\sigma}^+ = 0$. By (15), there exists a cell $M_\sigma \in \mathcal{N}_\sigma(K)$ and $\alpha_\sigma \in [0,1]$ such that

$$u_\sigma^n - u_K^n = \alpha_\sigma \left(u_{M_\sigma}^n - u_K^n\right), \text{ and therefore, } F_{K,\sigma}^- \left(u_\sigma^n - u_K^n\right) = \alpha_\sigma\, F_{K,\sigma}^- \left(u_{M_\sigma}^n - u_K^n\right).$$

Similarly, on the faces (including faces of $\mathcal{E}_{\rm N}$ and $\mathcal{E}_{\rm D}^+$) where the flow is leaving K, that is, $F_{K,\sigma}^+ \geqslant 0$, $F_{K,\sigma}^- = 0$, by (15), there exists $M_\sigma \in \mathcal{N}_\sigma(K)$ and $\alpha_\sigma \in [0,1]$ such that

$$u_\sigma^n - u_K^n = \alpha_\sigma \left(u_K^n - u_{M_\sigma}^n\right), \text{ and therefore } -F_{K,\sigma}^+ \left(u_\sigma^n - u_K^n\right) = \alpha_\sigma\, F_{K,\sigma}^+ \left(u_{M_\sigma}^n - u_K^n\right).$$

With these expressions, we thus obtain

$$\begin{aligned} \frac{|K|}{\delta t}\tilde{u}_K^{n+1} = &\left[\frac{|K|}{\delta t} - \sum_{\sigma \in \mathcal{E}(K)\setminus\mathcal{E}_{\rm D}} \alpha_\sigma\, |F_{K,\sigma}| - \sum_{\sigma \in \mathcal{E}(K)\cap\mathcal{E}_{\rm D}} F_{K,\sigma}^-\right] u_K^n \\ &+ \sum_{\sigma \in \mathcal{E}(K)\setminus\mathcal{E}_{\rm D}^-} \alpha_\sigma\, |F_{K,\sigma}|\, u_{M_\sigma}^n + \sum_{\sigma \in \mathcal{E}(K)\cap\mathcal{E}_{\rm D}^-} F_{K,\sigma}^-\, u_{{\rm D},\sigma}^n, \end{aligned}$$

which concludes the proof, because cfl $\leqslant 1$. □

For the sake of thoroughness, we now show that, under assumptions on the diffusion operator, this result yields a discrete maximum principle for the complete scheme (2). If we assume that both $f = 0$ and $g_{\rm N} = 0$, then the scheme (2) may be written as follows

$$\left((\mathrm{Id} - \kappa\delta t\ \Delta_{\mathcal{M}})\, u^{n+1}\right)_K = \tilde{u}_K^{n+1}, \tag{16}$$

DOI: 10.1002/fld

where $\tilde{u}_K^{n+1}$ is defined by (12) and $\mathrm{Id}-\kappa\delta t\,\Delta_{\mathcal{M}}$ stands for the operator acting on discrete functions, which maps $u=(u_K)_{K\in\mathcal{M}}$ to $(\mathrm{Id}-\kappa\delta t\,\Delta_{\mathcal{M}})\,u=(u_K-\kappa\delta t(\Delta_{\mathcal{M}}u)_K)_{K\in\mathcal{M}}$. Note that, from the definition of the discrete Laplace operator of Section 3.1, this operator is affine, and not linear, that is, in the case of non-homogeneous Dirichlet boundary conditions, $(\mathrm{Id}-\kappa\delta t\,\Delta_{\mathcal{M}})\,u$ does not vanish for $u=0$.

Theorem 3.2 (A discrete maximum principle)
Let us suppose that condition (15) holds, cfl $\leqslant 1$, and that both $f=0$ and $g_{\mathrm{N}}=0$. Then we have the following stability results:

(i) if $\kappa=0$, the solution to the scheme (2) satisfies a local maximum principle, namely $\forall K\in\mathcal{M}$, u_K^{n+1} is a convex combination of $\{u_K^n,(u_M^n)_{M\in\mathcal{N}_{\mathrm{m}}(K)},\ (u_{\mathrm{D},\sigma}^n)_{\sigma\in\mathcal{E}_{\mathrm{D}}^-\cap\mathcal{E}(K)}\}$.
(ii) if the discrete Laplace operator is such that
  (ii-a) for any constant function u such that $u\leqslant\min\{(u_{\mathrm{D},\sigma})_{\sigma\in\mathcal{E}_{\mathrm{D}}}\}$ (resp. $u\geqslant\max\{(u_{\mathrm{D},\sigma})_{\sigma\in\mathcal{E}_{\mathrm{D}}}\}$), $-\Delta_{\mathcal{M}}u\leqslant 0$ (resp. $-\Delta_{\mathcal{M}}u\geqslant 0$),
  (ii-b) all the entries of the inverse of the matrix $M_{I-\kappa\delta t\Delta}$ associated to the operator $\mathrm{Id}-\kappa\,\delta t\,\Delta_{\mathcal{M}}$ are non-negative, that is, $M_{I-\kappa\delta t\Delta}$ is a positive inverse matrix,
  the solution to the scheme (2) satisfies the following global maximum principle:

$$\forall K\in\mathcal{M},\quad \min\left\{(u_M^n)_{M\in\mathcal{M}},\ (u_{\mathrm{D},\sigma}^n)_{\sigma\in\mathcal{E}_{\mathrm{D}}^-},\ (u_{\mathrm{D},\sigma}^{n+1})_{\sigma\in\mathcal{E}_{\mathrm{D}}}\right\}\leqslant u_K^{n+1}$$
$$\leqslant\max\left\{(u_M^n)_{M\in\mathcal{M}},\ (u_{\mathrm{D},\sigma}^n)_{\sigma\in\mathcal{E}_{\mathrm{D}}^-},\ (u_{\mathrm{D},\sigma}^{n+1})_{\sigma\in\mathcal{E}_{\mathrm{D}}}\right\}. \tag{17}$$

Proof
Because we assume $f=0$ and $g_N=0$, when $\kappa=0$, the scheme (2) (or (16)) boils down to $\tilde{u}_K^{n+1}=u_K^{n+1},\ \forall K\in\mathcal{M}$, so item (i) is a straightforward consequence of Lemma 3.1. We now turn to item (ii) and define $\underline{u}$ by

$$\underline{u}=\min\ \left\{(u_K^n)_{K\in\mathcal{M}},\ (u_{\mathrm{D},\sigma}^n)_{\sigma\in\mathcal{E}_{\mathrm{D}}^-},\ (u_{\mathrm{D},\sigma}^{n+1})_{\sigma\in\mathcal{E}_{\mathrm{D}}}\right\}.$$

We have, $\forall K\in\mathcal{M}$:

$$\left((\mathrm{Id}-\kappa\delta t\ \Delta_{\mathcal{M}})\,(u^{n+1}-\underline{u})\right)_K=\tilde{u}_K^{n+1}-\underline{u}^n+\kappa\ \delta t\ (\Delta_{\mathcal{M}}\underline{u})_K. \tag{18}$$

By Lemma 3.1 and Assumption (ii-a), the right-hand side of this relation is non-negative. Because, by Assumption (ii-b), the operator at the left-hand side is associated to a positive inverse matrix, $u^{n+1}-\underline{u}$ is non-negative, which yields the first inequality of (17). The second inequality follows by the same computation with $\underline{u}=\max\left\{(u_K^n)_{K\in\mathcal{M}},\ (u_{\mathrm{D},\sigma}^n)_{\sigma\in\mathcal{E}_{\mathrm{D}}^-},\ (u_{\mathrm{D},\sigma}^{n+1})_{\sigma\in\mathcal{E}_{\mathrm{D}}}\right\}$. □

Remark 4
Assumptions (ii-a) and (ii-b) are satisfied with the usual two-point flux finite volume operator; unfortunately, this is not the case, in general, for the discrete Laplace operator introduced in Section 3.1 (neither, to our knowledge, by any linear discrete diffusion operator acting on general meshes).

3.2.2. A limitation procedure and the convection scheme. We now reformulate (15) to obtain a limitation procedure, that is, a constructive process to bound the face values $(u_\sigma^n)_{\sigma\in\mathcal{E}}$.

Let $\sigma\in\mathcal{E}_{\mathrm{int}}$; let us denote by V^- and V^+ the upstream and downstream cell separated by σ, and by $\mathcal{N}_\sigma(V^-)$ and $\mathcal{N}_\sigma(V^+)$ two sets of neighbouring cells of V^- and V^+, respectively. We would like to use (15), or possibly a stronger version of (15), to define an admissible interval for u_σ^n; this admissible interval must allow the usual upstream choice, that is, $u_\sigma^n=u_{V-}^n$. This is realized by the following two assumptions, provided that $V^-\in\mathcal{N}_\sigma(V^+)$ (Remark 5), which we thus suppose

(H1) – there exists $M\in\mathcal{N}_\sigma(V^+)$ such that $u_\sigma^n\in|[u_M^n,u_M^n+\dfrac{\zeta^+}{2}\left(u_{V^+}^n-u_M^n\right)]|$,

(H2) – there exists $M\in\mathcal{N}_\sigma(V^-)$ such that $u_\sigma^n\in|[u_{V-}^n,u_{V-}^n+\dfrac{\zeta^-}{2}\left(u_{V-}^n-u_M^n\right)]|$,

DOI: 10.1002/fld

where, for $a,\ b \in \mathbb{R}$, we denote by $|[a,b]|$ the interval $\{\alpha\, a + (1-\alpha)\, b,\ \alpha \in [0,1]\}$, and ζ^+ and ζ^- are two numerical parameters lying in the interval $[0,2]$. For $\sigma \in \mathcal{E}_{\rm N} \cup \mathcal{E}_{\rm D}^+$, condition (H1) is irrelevant, and the only acting constraint is condition (H2). The sets $\mathcal{N}_\sigma(V^-)$ and $\mathcal{N}_\sigma(V^+)$ have to be specified to complete the definition of the limitation process.

Let us show that (H1) and (H2) are sufficient conditions for (15) to hold. Let $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}(K)$. If $F_{K,\sigma} \leqslant 0$, that is K is the downstream cell for σ denoted earlier by V^+, because $\zeta^+ \in [0,2]$, condition (H1) yields that there exists $M \in \mathcal{M}$ such that $u_\sigma^n \in |[u_K^n, u_M^n]|$, which is (15). Otherwise, if $F_{K,\sigma} \geqslant 0$ and K is the upstream cell for σ denoted previously by V^-, condition (H2) yields that there exists $M \in \mathcal{M}$ such that $u_\sigma^n \in |[u_K^n, 2u_K^n - u_M^n]|$, so $u_\sigma^n - u_K^n \in |[0, u_K^n - u_M^n]|$, which is once again (15).

Remark 5
For $\sigma \in \mathcal{E}_{\rm int}$, because $V^- \in \mathcal{N}_\sigma(V^+)$, the upstream choice $u_\sigma^n = u_{V-}^n$ always satisfies the conditions (H1) and (H2), and is the only one to satisfy them if we choose $\zeta^- = \zeta^+ = 0$.

Remark 6 (One-dimensional case)
Let us take the example of an interface σ separating K_i and K_{i+1} in a 1D case (see Figure 3 for the notations), with a uniform meshing and a positive advection velocity, so that $V^- = K_i$ and $V^+ = K_{i+1}$. In 1D, a natural choice is $\mathcal{N}_\sigma(K_i) = \{K_{i-1}\}$ and $\mathcal{N}_\sigma(K_{i+1}) = \{K_i\}$.
In Figure 3, we sketch on the left, the admissible interval given by (H1) with $\zeta^+ = 1$ (green) and $\zeta^+ = 2$ (orange); on the right, the admissible interval given by (H2) with $\zeta^- = 1$ (green) and $\zeta^- = 2$ (orange). The parameters ζ^- and ζ^+ may be seen as limiting the admissible slope between $(\boldsymbol{x}_i, u_i^n)$ and $(\boldsymbol{x}_\sigma, u_\sigma^n)$ (with $\boldsymbol{x}_i$ the abscissa of the mass centre of K_i and $\boldsymbol{x}_\sigma$ the abscissa of σ), with respect to a left and right slope, respectively. For $\zeta^- = \zeta^+ = 1$, one recognises the usual minmod limiter (e.g. [16, Chapter III]).
Note that because in the example depicted in Figure 3, the discrete function u^n has an extremum in K_i, the combination of the conditions (H1) and (H2) imposes that, as usual, the only admissible value for u_σ^n is the upwind one.

We are now in a position to give the algorithm used for the discretization of the convection term:

1. Compute a tentative value $\tilde{u}_\sigma$ for the unknown at the face σ, by Relation (6), which yields an affine interpolation at the mass centre of the face.
2. For each face σ of the mesh, determine V^- and V^+ according to the sign of the mass flux through σ, and exploit (H1) and (H2) to obtain an admissible interval I_σ for the value of the unknown at the face, which, as explained in Remark 5, is not empty.
   This step depends on the definition of the sets $\mathcal{N}_\sigma(V^-)$ and $\mathcal{N}_\sigma(V^+)$. Here, we simply set $\mathcal{N}_\sigma(V^+) = \{V^-\}$; note however, that this choice prevents second order, because an affine function is not represented exactly (Remark 8). Two different choices of $\mathcal{N}_\sigma(V^-)$ are implemented (Figure 4):
   (a) $\mathcal{N}_\sigma(V^-)$ is defined as the set of 'upstream cells' to V^-, that is $\mathcal{N}_\sigma(V^-) = \{L \in \mathcal{M},\ L$ shares a face σ with V^- and $F_{V^-,\sigma} < 0\}$,

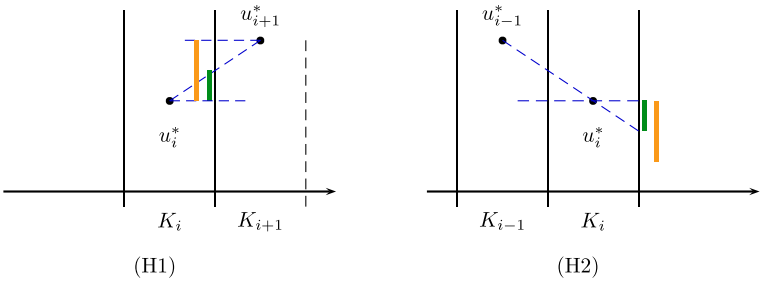
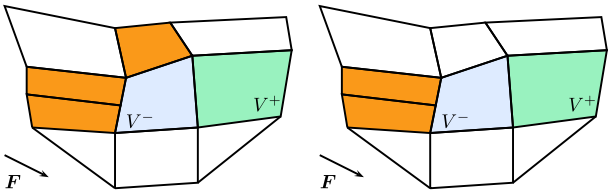


Figure 3. Conditions (H1) and (H2) in 1D.

DOI: 10.1002/fld



(a) (b)

Figure 4. Notations for the definition of the limitation process. In orange, control volumes of the set $\mathcal{N}_\sigma(V^-)$ for $\sigma = V^-|V^+$, with a constant advection field $\boldsymbol{F}$: upwind cells (a) or opposite cells (b).

(b) when this makes sense, that is, with a mesh obtained by Q_1 mappings from the $(0,1)^d$ reference element), $\mathcal{N}_\sigma(V^-)$ may be chosen as the opposite cells to σ in V^-. Note that, for a structured mesh, this choice allows to recover the usual minmod limiter.

3. Compute u_σ as the nearest point to $\tilde{u}_\sigma$ in I_σ.

Remark 7 (Interpolation of the unknown at the face)
The tentative (i.e. before limitation) value of the unknown at the faces is given by the same interpolation as for the diffusion (SUSHI) scheme, and so takes the form, for the non-Dirichlet faces:

$$\forall \sigma \in \mathcal{E}_{\text{int}} \cup \mathcal{E}_{\text{N}}, \qquad \tilde{w}_\sigma = \sum_{K \in \mathcal{M}} \beta_\sigma^K w_K. \tag{19}$$

The computation of the coefficients (β_σ^K) is performed as follows:

- We first consider several possible families $(\beta_\sigma^M)_{M\in\mathcal{M}}$ such that (19) holds: for an internal face $\sigma = K|L$, we consider all the families $(\beta_\sigma^M)_{M\in\mathcal{M}}$, which satisfy (19) and are such that $\beta_\sigma^M = 0$ except for $M = K$, $M = L$, and for one (in 2D) or two (in 3D) cell(s) M, which share a face with K or L; for an external face of a cell K, we consider all the families $(\beta_\sigma^M)_{M\in\mathcal{M}}$ which satisfy (19) and are such that $\beta_\sigma^M = 0$ except for $M = K$ and for two (in 2D) or three (in 3D) other cells M sharing a face with K.
- Then we have to choose among the obtained families. We first choose among the families that yield a convex combination in (19) (i.e. which satisfy $\beta_\sigma^K \geqslant 0$, $\forall K \in \mathcal{M}$), if any. If, for one of these convex combinations, only two coefficients differ from zero (which means that the centre of mass of the face $\boldsymbol{x}_\sigma$ is aligned with the centroids of two cells), then it is chosen for the computations. Otherwise, for each combination, we compute the real number $\beta = \max_{\beta_\sigma^K \neq 0} |\beta_\sigma^K - 0.5|$ and choose the combination that leads to the minimum value for β; loosely speaking, we thus pick the configuration where $\boldsymbol{x}_\sigma$ is best located 'at the centre' of the convex set. If there is no convex combination, we turn to non-convex ones (which is almost always the case for an external face), and choose once again the one that is characterized by the lowest parameter β.

Remark 8 (Reconstruction of affine functions)
Let us suppose that we are trying to transport (i.e. in fact, to keep constant) in $\mathbb{R}^2$ the initial function $u(\boldsymbol{x}) = \boldsymbol{x}_2$, with a constant advection velocity $\boldsymbol{v} = (1,0)^t$. Let K_1 and K_2 be two cells, of mass centre located at $\boldsymbol{x}_1 = (1,0)^t$ and $\boldsymbol{x}_2 = (2,0)^t$, respectively, and suppose that we initialize the scheme by setting for any $K \in \mathcal{M}$ the value u_K at the mean value of $u(\boldsymbol{x})$, so that $u_{K_1} = u_{K_2} = 0$. Let the face $\sigma = K_1|K_2$ be vertical, with a mass centre located at $\boldsymbol{x}_\sigma = (1.5, 0.5)^t$. Any affine reconstruction must yield $u_\sigma = u(\boldsymbol{x}_\sigma) = 0.5$ for the value u_σ of the approximation of u on σ, but condition (H1) with $\mathcal{N}_\sigma(K_2) = \{K_1\}$ yields $u_\sigma = 0$.

DOI: 10.1002/fld

4. NUMERICAL TESTS

The computations performed in this section were performed with the open-source ISIS computer code [17], developed at IRSN on the basis of the software component library PELICANS [18].

4.1. Transport of smooth functions

We first assess the accuracy of the scheme for transport problems of smooth functions. We begin with a (apparently very simple) steady test where the solution is given by $u = \boldsymbol{x}_2$ and the velocity is $\boldsymbol{v} = (1, 0)^t$. The domain is $\Omega = (0, 1)^2$, and the (quadrangular) mesh is obtained by first building a structured square grid of step h, and then applying a random displacement of length $0.2\,h$ to each node located in the subdomain $(0.1, 0.9)^2$. A similar mesh (but with a displacement of $0.3h$ all the nodes of the domain) is sketched in Figure 10.

As explained in Remark 8, avoiding a distortion of the solution necessitates to choose sets $\mathcal{N}_\sigma(V^-)$ and $\mathcal{N}_\sigma(V^+)$ large enough. We first check that, when all the neighbours of (i.e. the control volumes sharing a face with) V^+ are included in $\mathcal{N}_\sigma(V^+)$ and all the neighbours of V^- except V^+ are included in $\mathcal{N}_\sigma(V^-)$, the steady state solution is preserved by the scheme. Note that with a mesh unstructured up to the boundary, the preservation of the steady state would necessitate a specific treatment of the boundary conditions, taking into account the values taken by the solution on $\partial\Omega$ in the limitation process.

However, when performing numerical experiments with a non-trivial solution, choosing large sets $\mathcal{N}_\sigma(V^-)$ and $\mathcal{N}_\sigma(V^+)$ seems to yield a scheme which is not stable enough, in the sense that its solution presents (limited, because a local discrete maximum principle holds) oscillations. We now turn to one of the stricter limitations announced in the preceding section, where $\mathcal{N}_\sigma(V^+) = \{V^-\}$ and $\mathcal{N}_\sigma(V^-)$ only contains the opposite cell. The time step is given by $\delta t = h/10$ (so the CFL number is close to 0.35). At $t = 1$ (so when the 'numerical steady state' is reached), the norm of the difference between the numerical and exact solution, for various meshes, is as follows:

Initial mesh	20×20	40×40	80×80	160×160
$10^3\times$ error (L^1 norm)	2.59	1.43	0.73	0.368

This corresponds to a first-order convergence.

We then turn to an unsteady test, the so-called 'double-sine' wave simulated in [11]. The solution reads as follows:

$$u = \sin(2\pi \boldsymbol{x}_1)\ \sin(2\pi \boldsymbol{x}_2),$$

and the velocity is $\boldsymbol{v} = (2, 1)^t$. The computational domain is $\Omega = (0, 2.5) \times (0, 1.5)$. Initial and boundary conditions are prescribed so as to be consistent with the solution. Meshes are built from a structured grid by cutting each square cell in two subtriangles along the diagonal parallel to the $(1, 1)^t$ vector. The time step is given by $\delta t = h/25$ (so the CFL number is 0.32). We choose $\mathcal{N}_\sigma(V^+) = \{V^-\}$ and the set $\mathcal{N}_\sigma(V^-)$ such that it contains the upstream cells.

We show in Figure 5 the value of the solution obtained with the parameters $\zeta^+ = \zeta^- = 1$ at $t = 1$ along the segment of starting and ending points $(2.25, 1)^t$ and $(2.25, 1.5)^t$, respectively (so a structure which results from the transport of a bump initially present in the domain). The maximal value along this curve is as follows:

Initial mesh	$h = 1/20$	$h = 1/40$	$h = 1/80$
Maximal value	0.69	0.85	0.93

These results seem to indicate that the present scheme is less diffusive than the classical limitation algorithms tested in [11], and more than the new algorithm proposed in this latter work.

DOI: 10.1002/fld

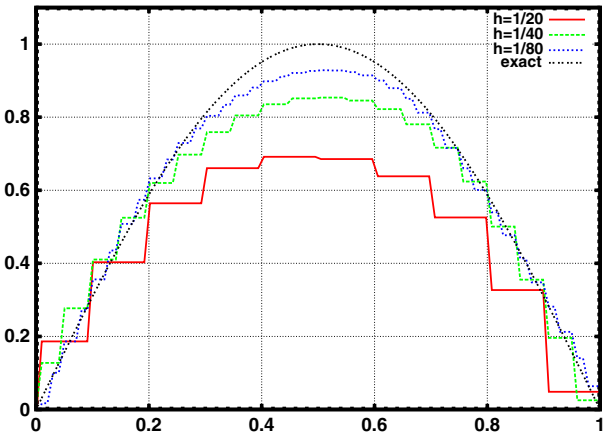


Figure 5. Double sine wave - Solution obtained with various meshes along the line joining the point $(2.25, 1)^t$ to $(2.25, 1.5)^t$, and analytical solution – $\zeta^+ = \zeta^- = 1$.

Choosing now $\zeta^+ = \zeta^- = 2$ strongly reduces the diffusion (especially for the coarsest mesh) because the same peak values (Figure 6) become:

Initial mesh	$h = 1/20$	$h = 1/40$	$h = 1/80$
Maximal value	0.90	0.95	0.982

However, for the finest mesh, oscillations along the flow directions appear at the sides of the sinusoidal bump, in the areas where the isovalues of the solution are tangent to the flow, so we cannot recommend these values for ζ^+ and ζ^- as standard parameters for the scheme.

4.2. Transport of irregular functions

We now address the pure transport (i.e. without diffusion) of an irregular function defined on $\Omega = (-1, 1)^2$ as follows:

$$\begin{aligned} \text{for } \boldsymbol{x} \in (0.1, 0.6) \times (-0.25, 0.25), \quad & u = 1, \\ \text{if } r < 0.35, \quad & u = 1 - r/0.35, \\ \text{otherwise}, \quad & u = 0. \end{aligned}$$

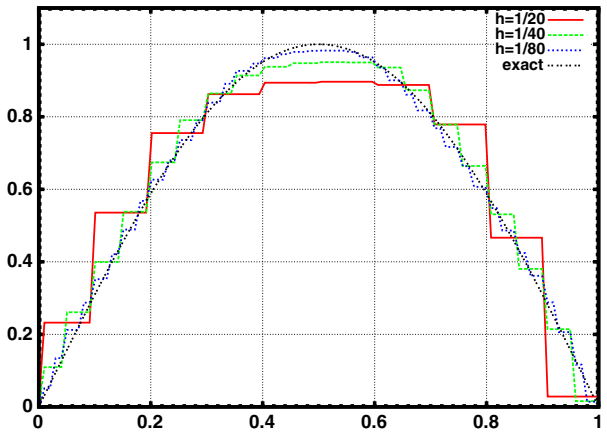


Figure 6. Double sine wave - Solution obtained with various meshes along the line joining the points $(2.25, 1)^t$ to $(2.25, 1.5)^t$, and analytical solution–$\zeta^+ = \zeta^- = 2$.

DOI: 10.1002/fld

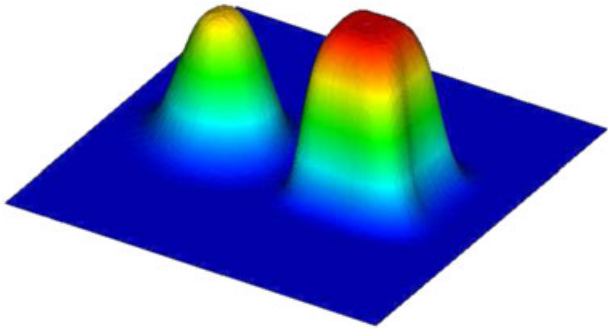


Figure 7. Transport of irregular functions - Results obtained with the usual MUSCL scheme and a minmod limiter, for a uniform 120×120 Cartesian grid.

where r stands for the distance from the current point $\boldsymbol{x}$ to the point $(-0.45, 0)^t$, that is, $r^2 = (\boldsymbol{x}_1 + 0.45)^2 + \boldsymbol{x}_2^2$. The advection field is given by $\boldsymbol{v}(\boldsymbol{x}) = 2\,(\boldsymbol{x}_2, -\boldsymbol{x}_1)^t$, so, at the end of n complete revolutions, the solution is identical to the initial condition for $t = n\pi$, $n \in \mathbb{N}$.

We begin with uniform Cartesian grids. If we set $\mathcal{N}_\sigma(V^-)$ to the opposite cell and choose $\zeta^+ = \zeta^- = 1$, the proposed scheme boils down to a usual MUSCL scheme with a minmod limiter (Remark 6). Results obtained with a 120×120 mesh and $\delta t = \pi/(125 * n)$ (which yields a cfl number close to 1) at $t = \pi$ are plotted in Figure 7.

In Figure 8, we plot the value of the unknown along the $x-$axis, for the same mesh and time step and various options of the scheme: the MUSCL minmod one, the upwind scheme (obtained here by taking $\zeta^+ = \zeta^- = 0$), and two variants with less stringent limitations obtained by taking $\zeta^+ = \zeta^- = 2$ and by enlarging $\mathcal{N}_\sigma(V^-)$ to the set of upstream cells. At first glance, results may seem better with less limitation, but the shape of the initial condition is deformed, as may be seen in Figure 9.

We next turn to unstructured meshes. Starting from a regular Cartesian grid and applying a random displacement of length $0.3h$ to each node, we first obtain an unstructured quadrangular mesh; then, splitting each cell in four along its diagonals, we obtain a simplicial mesh. The coarsest meshes used in this study, together with the discrete initial condition (determined for each cell K as the value of the initial data at $\boldsymbol{x}_K$), are plotted in Figure 10. From now on, we restrict the choice for $\mathcal{N}_\sigma(V^-)$ to the opposite mesh for quadrangles and to the set of upstream cells for simplices, and we set $\zeta^+ = \zeta^- = 1$.

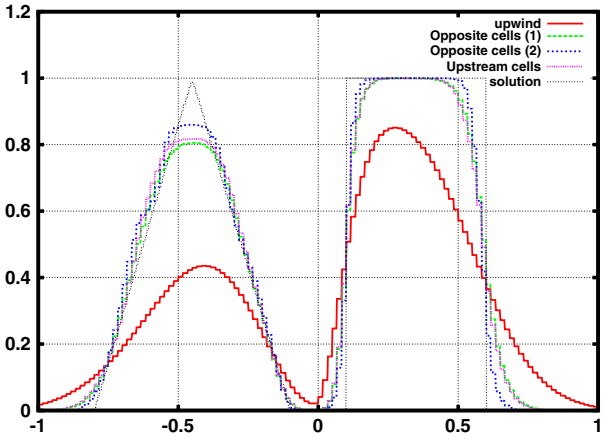


Figure 8. Transport of irregular functions - Value of the unknown along the $x-$axis obtained with the usual upwind scheme, the MUSCL minmod scheme (curve Opposite cell (1)), the same scheme with $\zeta^+ = \zeta^- = 2$ (curve Opposite cell (2)), and taking for $\mathcal{N}_\sigma(V^-)$ the set of upstream cells, for a uniform 120×120 Cartesian grid.

DOI: 10.1002/fld

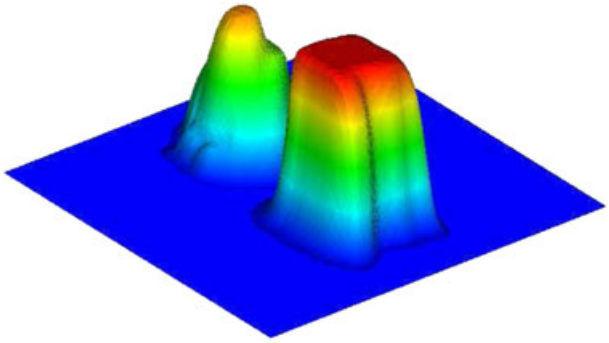


Figure 9. Transport of irregular functions - Results obtained with less limitation ($\zeta^+ = \zeta^- = 2$), for a uniform 120×120 Cartesian grid.

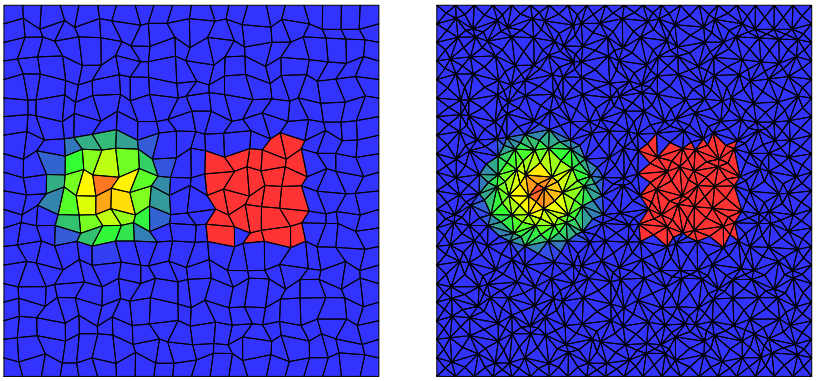


Figure 10. Transport of irregular functions - Mesh and initial value for quadrangular (left) and simplicial (right) meshes obtained from an initial 20×20 structured Cartesian grid.

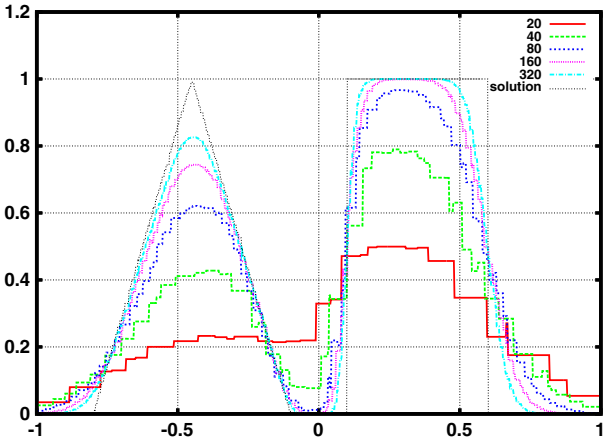


Figure 11. Transport of irregular functions - Value of the unknown along the $x-$axis, obtained with the quadrangular cells, as a function of the mesh step.

The value of the unknown along the $x-$axis is plotted for meshes obtained from $n \times n$ structured grids, with $n = 20$, 40, 80, 160 and 320, and $\delta t = \pi/(175 * n)$, for quadrangles (Figure 11) and $n = 20$, 40, 80 and 160, and $\delta t = \pi/(600 * n)$, for triangles (Figure 12). In both cases, the cfl number is near to 1, and is the same for all the computations performed with the same family of meshes (i.e. quadrangular or simplicial meshes). The main qualitative effect of using an unstructured mesh seems to be an additional smearing of the solution.

DOI: 10.1002/fld

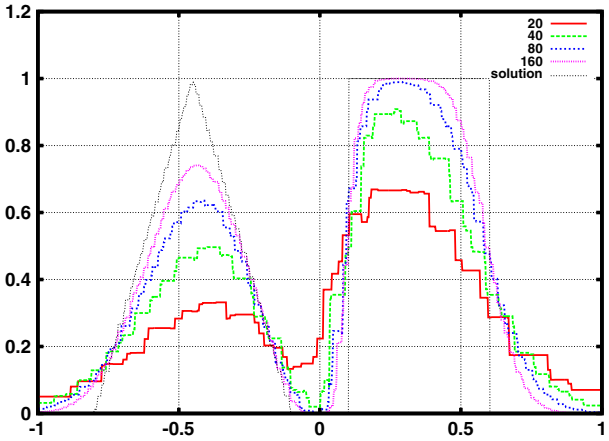


Figure 12. Transport of irregular functions - Value of the unknown along the $x-$axis, obtained with simplicial cells, as a function of the mesh step.

The difference between the obtained solution, in a discrete L^1-norm defined by

$$\| u \|_{L^1,\mathcal{M}} = \sum_{K \in \mathcal{M}} |K| \, |u(\boldsymbol{x}_K)|,$$

is given in the following table, as a function of the initial regular grid, for the different computations already invoked in this study.

Initial mesh	20×20	40×40	80×80	160×160	320×320
Structured mesh	0.38	0.21	0.12	0.077	0.042
Quadrangles	0.42	0.28	0.18	0.12	0.081
Triangles	0.37	0.26	0.19	0.13	//

As may be expected, the accuracy is lower with unstructured meshes, and the fact that the numerical diffusion is greater is confirmed by the comparison of the trace of the solutions along the line $\boldsymbol{x}_2 = 0$ obtained with the various meshes built in this study from the 160×160 grid, displayed in Figure 13.

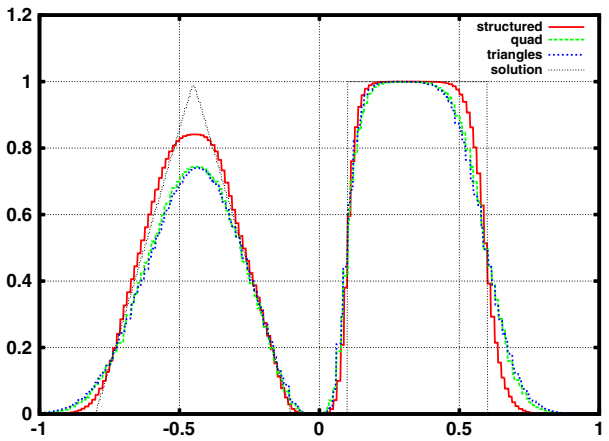


Figure 13. Transport of irregular functions - Value of the unknown along the $x-$axis, obtained with various meshes, built from a regular 160×160 mesh.

DOI: 10.1002/fld

To conclude, we assess the capability of the scheme to deal with a locally refined non-conforming mesh. To this purpose, we start from a regular quadrangular mesh and split in four the cells located above the line $\boldsymbol{x}_2 = 0$. Results obtained at $t = \pi$, starting with a 80×80 regular grid, are displayed in Figures 14 and 15. No spurious numerical phenomenon is observed (especially near hanging nodes), and the computation performed with the partially refined mesh appears less diffusive. As predicted by the theory, here as in all the performed test cases, no overshoot or undershoot of the solution is observed (i.e. here, the solution remains in the interval $[0, 1]$).

4.3. A convection–diffusion case

We now turn to a convection–diffusion tests case, built by combining a classical solution of the heat equation with a constant skew-to-the mesh transport. The computational domain is $\Omega = (0, 2) \times (0, 2)$, the advection velocity is $\boldsymbol{v} = (0.8, 0.8)^t$, the solution is given by

$$u = \frac{1}{4t+1} \exp\left(\frac{X^2 + Y^2}{\kappa(4t+1)}\right), \qquad \begin{bmatrix} X \\ Y \end{bmatrix} = \boldsymbol{x} - \left(\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} + t\boldsymbol{v}\right),$$

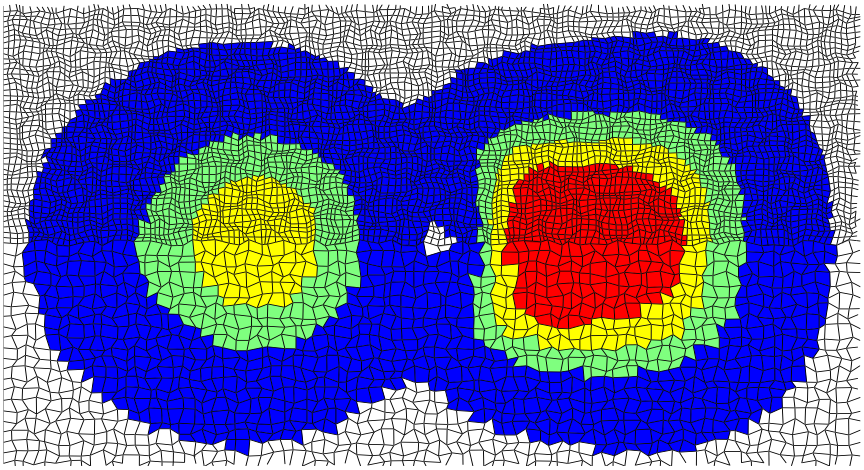


Figure 14. Transport of irregular functions - Solution at $t = \pi$ obtained with a locally refined mesh (part of the computational domain only). If $u \leqslant 0.01$, meshes are coloured in white; if $0.01 < u \leqslant 0.25$, in blue; if $0.25 < u \leqslant 0.5$, in green; if $0.5 < u \leqslant 0.75$, in yellow; If $0.75 < u$, in red.

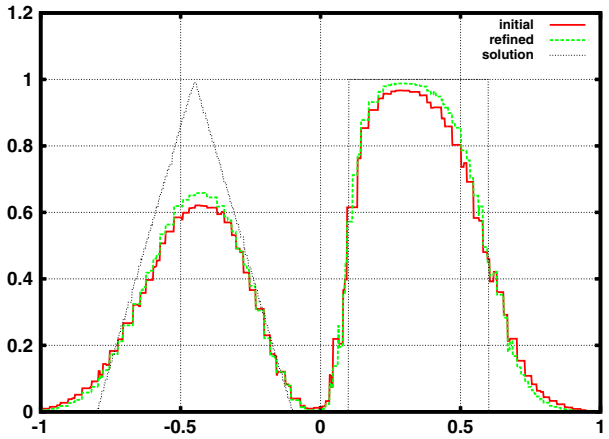


Figure 15. Transport of irregular functions - Value of the unknown along the $x-$axis, obtained with a locally refined mesh and with the quadrangular initial (i.e. before refinement) mesh.

DOI: 10.1002/fld

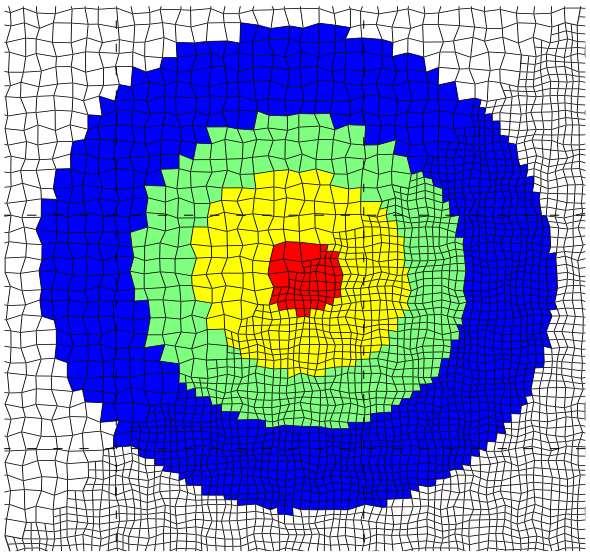


Figure 16. Convection–diffusion case - Solution obtained at $t = 1.2$ with a locally refined mesh (zoom on a part of the computational domain). If $u \leqslant 0.01$, meshes are coloured in white; if $0.01 < u \leqslant 0.05$, in blue; if $0.05 < u \leqslant 0.1$, in green; if $0.1 < u \leqslant 0.15$, in yellow; if $0.15 < u$, in red.

with $\kappa = 0.01$. The function u satisfies the advection–diffusion equation (1a) with $f = 0$, and initial and Dirichlet boundary conditions given by value of u at $t = 0$ and on $\partial\Omega$, respectively.

As in the previous section, we use meshes of quadrangles obtained by perturbation of regular grids, by a displacement of each node in a random direction, here of length $0.2h$. We reduce $\mathcal{N}_\sigma(V^-)$ to the opposite mesh of σ in V^- and choose $\zeta^+ = \zeta^- = 1$.

The L^2-norm of the difference between the numerical and continuous solution at $t = 1.2$ is given for several meshes (with a time step adjusted accordingly to have cfl ≈ 0.5) in the following table. The convergence seems to be rather fast for coarse meshes, then slows down, the convergence rate

Initial mesh	40×40	80×80	160×160	320×320
Time step	0.005	0.0025	0.001	0.0005
Error (L^2-norm)	0.0095	0.0028	0.00095	0.00042

however remaining greater than 1.

We now assess the capability of the scheme to work on a locally refined mesh. We start from the mesh of quadrangles obtained from the 80×80 mesh and cut in four sub-quadrangles the meshes located under the line $\boldsymbol{x}_2 = \boldsymbol{x}_1$. The result obtained at $t = 1.2$ in the upper left part of the computational domain (the part where the solution varies at that time) is plotted in Figure 16. We observe the absence of numerical perturbations at the edges where the mesh is non-conforming. In the upper part of the domain (i.e. $\boldsymbol{x}_2 > \boldsymbol{x}_1$), the numerical diffusion appears slightly larger, which is consistent with the fact that the mesh is coarser.

5. CONCLUSION

In this paper, we described a finite volume scheme for the solution of the advection–diffusion equation, which copes with almost arbitrary meshes. This scheme combines two ingredients:

- a discrete diffusion operator, which is both consistent and stable,
- a non-linear discrete transport operator using a prediction/limitation procedure, with a limitation step that ensures the satisfaction of a local maximum principle without invoking any geometrical argument.

DOI: 10.1002/fld

The material presented here may be developed in several directions:

- First, the proposed limitation procedure may be used to complement any other existing algorithm, as a final step to ensure the local maximum principle without any restriction on the mesh; doing so, the parameters should probably be tuned to limit as less as possible (i.e. to enlarge as much as possible the admissible interval for the value at the face),
- Second, we paid no particular attention here to the reconstruction of the value at the face, and, especially for the transport of smooth functions, it is probably possible to design a more accurate evaluation, for instance using a least squares technique,
- Last, but not least, still for the transport of smooth functions, it is certainly preferable to switch to a second-order in time scheme.

In addition, the convection scheme presented here extends to variable density flows, that is, to a balance equation for $\bar{u}$ of the form $\partial_t(\varrho\bar{u}) + \mathrm{div}(\varrho\bar{u}\boldsymbol{v}) - \mathrm{div}(\kappa\boldsymbol{\nabla}\bar{u}) = f$, where the density ϱ and the velocity field $\boldsymbol{v}$ are linked by the usual mass balance equation $\partial_t\varrho + \mathrm{div}(\varrho\boldsymbol{v}) = 0$.

REFERENCES

1. Eymard R, Herbin R. A new colocated finite volume scheme for the incompressible Navier–Stokes equations on general non matching grids. *Comptes Rendus Mathématique. Académie des Sciences. Paris* 2007; **344**:659–662.
2. Agelas L, Di Pietro DA. Benchmark on Anisotropic Problems—a symmetric finite volume scheme for anisotropic heterogeneous second-order elliptic problems. In *FVCA5 – Finite Volumes for Complex Applications V*. John Wiley & Sons: Hoboken, USA, 2008; 705–716.
3. Eymard R, Gallouët T, Herbin R. Discretisation of heterogeneous and anisotropic diffusion problems on general non-conforming meshes—SUSHI: a scheme using stabization and hybrid interfaces. *IMA Journal of Numerical Analysis* 2010; **30**:1009–1043.
4. Harten A. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics* 1983; **49**:357–393.
5. Sweby PK. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis* 1984; **21**:995–1011.
6. Van Leer B. Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second-order scheme. *Journal of Computational Physics* 1974; **14**:361–370.
7. Barth T, Ohlberger M. Finite volume methods: foundation and analysis. In *Encyclopedia of Computational Mechanics, Volume 1, Chapter 15*. John Wiley & Sons: Hoboken, USA, 2004; 439–474.
8. Buffard T, Clain S. Monoslope and multislope MUSCL methods for unstructured meshes. *Journal of Computational Physics* 2010; **229**:3745–3776.
9. Calgaro C, Chane-Kane E, Creusé E, Goudon T. L^∞-stability of vertex-based MUSCL finite volume schemes on unstructured grids: simulation of incompressible flows with high density ratios. *Journal of Computational Physics* 2010; **229**:6027–6046.
10. Manzini G, Russo A. A finite volume method for advection–diffusion problems in convection-dominated regimes. *Computer Methods in Applied Mechanics and Engineering* 2008; **197**:1242–1261.
11. Park JS, Yoon S-H, Kim C. Multi-dimensional limiting process for hyperbolic conservation laws on unstructured grids. *Journal of Computational Physics* 2010; **229**:788–812.
12. Clain S. Finite volume L^∞-stability for hyperbolic scalar problems. *submitted* 2010.
13. Clain S, Clauzon V. L^∞ stability of the MUSCL methods. *Numerische Mathematik* 2010; **116**:31–64.
14. Tran QH. A scheme for multi-dimensional linear advection with accuracy enhancement based on a genuinely one-dimensional min–max principle. *In preparation* 2010.
15. Chénier E, Eymard R, Herbin R. A collocated finite volume scheme to solve free convection for general non-conforming grids. *Journal of Computational Physics* 2009; **228**:2296–2311. to appear, see also http://hal.archives-ouvertes.fr.
16. Godlewski E, Raviart P-A. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Number 118 in Applied Mathematical Sciences. Springer: New York, 1996.
17. ISIS. A CFD computer code for the simulation of reactive turbulent flows. https://gforge.irsn.fr/gf/project/isis.
18. PELICANS. Collaborative development environment. https://gforge.irsn.fr/gf/project/pelicans.

DOI: 10.1002/fld