

Let  $\Omega \subset \mathbb{R}^3$  be a bounded Lipschitz domain of our interest, and  $\partial\Omega$  its boundary. In this work, our aim is to describe a numerical approximation of the scalar conservation law, also known as the continuity equation

$$\phi_t + \nabla \cdot (\phi \mathbf{u}) = 0, \quad (1)$$

where  $\phi = \phi(\mathbf{x}, t)$  is the conserved scalar quantity,  $(\mathbf{x}, t) \in \Omega \times [0, T]$ . For simplicity, we consider velocity fields  $\mathbf{u} = \mathbf{u}(\mathbf{x})$  that are functions of the position only. Initial and Dirichlet boundary conditions are given

$$\phi(\mathbf{x}, 0) = \phi^0(\mathbf{x}), \quad \mathbf{x} \in \Omega \quad (2)$$

$$\phi(\mathbf{x}_b, t) = \phi_b(t), \quad (\mathbf{x}, t) \in \Omega \times [0, T]. \quad (3)$$

### 0.0.1 Discretization of the domain

In case of a complex computational domain, it is often necessary to use unstructured polyhedral mesh. The finite volume method (FVM) is well suited for space-discretization of arbitrary shape. In FVM, the unknowns are the cell averages. A particular finite volume scheme is defined by its approximation of the fluxes at cell faces. Let  $p$  be a finite volume cell. We discretize the domain  $\Omega$  by open non-overlapping polyhedral sets  $\Omega_p \subset \Omega$ . Let  $\bar{\Omega} = \bigcup_{p \in \mathcal{I}} \bar{\Omega}_p \subset \mathbb{R}^3$  be a computational domain, where an open set  $\Omega_p$  with the non-zero volume  $|\Omega_p|$  is a cell in discretized domain and  $\mathcal{I}$  is the set of cell indices,  $p \in \mathcal{I}$ .

A distorted cell means that the cell has a nonplanar face, which usually exists when there is a complex computation boundary. It is an unrealistic assumption that all faces of  $|\Omega_p|$  for all  $p \in \mathcal{I}$  are of a planar shape in a case of industrial problems.

Whenever a face is not triangle, the face is tessellated into triangles; from a given center of the face, two consecutive vertices of the face are selected to construct one triangle. A center  $\mathbf{x}^*$  of face whose vertices are  $\mathbf{x}_i, i = 1, \dots, r$  is usually given by the area average. To make it simple, using a cyclic notation  $\mathbf{x}_{r+1} = \mathbf{x}_1$ , the area average face center is computed by

$$\mathbf{x}^* = \frac{\sum_{i=1}^r |\Delta_i| \bar{\mathbf{x}}_i}{\sum_{i=1}^r |\Delta_i|} \quad (4)$$

where  $|\Delta_i|$  is the area of  $\Delta_i$ ,  $\bar{\mathbf{x}}_i$  is the center of mass on  $\Delta_i$ , and  $\Delta_i$  is a triangle of  $\mathbf{x}_i, \mathbf{x}_{i+1}$ , and  $\mathbf{x}_0 = \frac{1}{r} \sum_{j=1}^r \mathbf{x}_j$  for  $i = 1, \dots, r$ . Then, the  $i$ -th tessellated triangle at the

face is defined by three points:  $\mathbf{x}_i, \mathbf{x}_{i+1}$ , and  $\mathbf{x}^*$ . The index set  $\mathcal{F}$  denotes all triangles  $e_f, f \in \mathcal{F}$  tessellated from a face or a triangle face of  $\Omega_p$  for all  $p \in \mathcal{I}$ . We denote  $\mathbf{x}_f$  as the center of triangle  $e_f$ .

To indicate the neighbor cells of  $\Omega_p$  we consider only the cells whose face is shared:

$$\mathcal{N}_p = \{q \in \mathcal{I} : \text{there exists a face } e_f \subset \partial\Omega_q \cap \partial\Omega_p, f \in \mathcal{F}\}.$$

The faces of  $\Omega_p$  are indicated by two sets:

$$\mathcal{F}_p = \{f \in \mathcal{F} : e_f \subset \partial\Omega_p\} \quad \text{and} \quad \mathcal{B}_p = \{f \in \mathcal{F} : e_f \subset \partial\Omega_p \cap \partial\Omega\}.$$

If  $\Omega_p$  is a cell whose all faces are not overlapped to  $\partial\Omega$ , then we call the cell as an internal cell and  $\mathcal{B}_p = \emptyset$ . Otherwise, we call the cell as a boundary cell. Similarly, an internal face  $e_f, f \in \mathcal{F}$ , means that the face is not overlapped to  $\partial\Omega$  and a boundary face  $e_b, b \in \mathcal{F}$ , is a part of  $\partial\Omega$ . Throughout the rest of paper, the subscript  $b$  indicates the face index whose face is a boundary face. When a quantity is defined on a face  $e_f$ , and it depends on a cell sharing the face, the cell index should be explicitly indicated at the first subscript and the face index at the second subscript. For instance, for a face  $e_f, f \in \mathcal{F}_p$ , the outward normal vector to the face is indicated by  $\mathbf{n}_{pf}$ . Note that we also use the length of the normal vector as the area of the face, that is,  $|\mathbf{n}_{pf}| = |e_f|$ . If a face  $e_f$  is an internal face for  $f \in \mathcal{F}_p$ , there exists a cell  $\Omega_q, q \in \mathcal{I}$ , such that  $e_f \subset \partial\Omega_p \cap \partial\Omega_q$ . Then, clearly,  $\mathbf{n}_{pf} = -\mathbf{n}_{qf}$ . Whenever a directional vector is denoted, we use a notation  $\mathbf{d}$  with relevant indices:

$$\mathbf{d}_{pq} = \mathbf{x}_q - \mathbf{x}_p, \quad p, q \in \mathcal{I}.$$

## 0.0.2 Numerical scheme

We integrate (1) in space and time

$$\int_{t^{n-1}}^{t^n} \int_{\Omega_p} \phi_t + \int_{t^{n-1}}^{t^n} \int_{\Omega_p} \nabla \cdot (\phi \mathbf{u}) = 0. \quad (5)$$

In the first term we change the order of integration and apply the Newton-Leibniz theorem for the time integration

$$\int_{t^{n-1}}^{t^n} \int_{\Omega_p} \phi_t = \int_{\Omega_p} \int_{t^{n-1}}^{t^n} \phi_t = \int_{\Omega_p} (\phi^n(\mathbf{x}) - \phi^{n-1}(\mathbf{x})) = |\Omega_p|(\bar{\phi}_p^n - \bar{\phi}_p^{n-1}),$$

where we denote  $\phi(\mathbf{x}, t^n) = \phi^n(\mathbf{x})$ , and,  $\bar{\phi}_p^n$  is the cell average at time  $t^n$  such that

$$\bar{\phi}_p^n = \frac{1}{|\Omega_p|} \int_{\Omega_p} \phi^n(\mathbf{x}). \quad (6)$$

Next we apply the divergence theorem to the volume integral

$$\int_{\Omega_p} \nabla \cdot (\phi \mathbf{u}) = \int_{\partial\Omega_p} (\phi \mathbf{u}) \cdot \frac{\mathbf{n}_p}{|\mathbf{n}_p|} = \sum_{f \in \mathcal{F}_p} \int_{e_f} (\phi \mathbf{u}) \cdot \frac{\mathbf{n}_{pf}}{|\mathbf{n}_{pf}|},$$

where  $\mathbf{n}_{pf}$  is an outward facing normal vector of the face  $e_f$ . Thus, we can rewrite the integral form of the conservation law (5) as

$$|\Omega_p|(\bar{\phi}_p^n - \bar{\phi}_p^{n-1}) + \int_{t^{n-1}}^{t^n} \sum_{f \in \mathcal{F}_p} \int_{e_f} (\phi \mathbf{u}) \cdot \frac{\mathbf{n}_{pf}}{|\mathbf{n}_{pf}|} = 0, \quad (7)$$

which is still exact for a finite volume cell  $p$ .

Throughout this text, we will approximate the cell average with the cell center value

$$\bar{\phi}_p^n \approx \phi_p^n.$$

In order to show that this is a second order approximation, we describe the spatial variation of  $\phi$  within an element via Taylor series expansion around the cell center  $\mathbf{x}_p$  as

$$\phi^n(\mathbf{x}) = \phi_p^n + \nabla \phi_p^n \cdot (\mathbf{x} - \mathbf{x}_p) + \mathcal{O}(|\mathbf{x} - \mathbf{x}_p|^2). \quad (8)$$

We substitute (8) to (6)

$$\begin{aligned} \bar{\phi}_p^n &= \frac{1}{|\Omega_p|} \int_{\Omega_p} \phi^n(\mathbf{x}) \\ &= \frac{1}{|\Omega_p|} \int_{\Omega_p} [\phi_p^n + \nabla \phi_p^n \cdot (\mathbf{x} - \mathbf{x}_p) + \mathcal{O}(|\mathbf{x} - \mathbf{x}_p|^2)] \\ &= \frac{1}{|\Omega_p|} \int_{\Omega_p} \phi_p^n + \frac{1}{|\Omega_p|} \int_{\Omega_p} \nabla \phi_p^n \cdot (\mathbf{x} - \mathbf{x}_p) + \frac{1}{|\Omega_p|} \int_{\Omega_p} \mathcal{O}(|\mathbf{x} - \mathbf{x}_p|^2) \\ &= \phi_p^n + \frac{1}{|\Omega_p|} \int_{\Omega_p} \mathcal{O}(|\mathbf{x} - \mathbf{x}_p|^2). \end{aligned} \quad (9)$$

The second integral equals to 0, since  $\mathbf{x}_p$  is the cell center. For more details, see, e.g., [5, 3].

We approximate the expression under the integral with the face center values

$$\int_{e_f} (\phi \mathbf{u}) \cdot \frac{\mathbf{n}_{pf}}{|\mathbf{n}_{pf}|} \approx \int_{e_f} \phi(\mathbf{x}_{pf}, t) \mathbf{u}(\mathbf{x}_{pf}) \cdot \frac{\mathbf{n}_{pf}}{|\mathbf{n}_{pf}|} = \phi(\mathbf{x}_{pf}, t) \mathbf{u}(\mathbf{x}_{pf}) \cdot \mathbf{n}_{pf},$$

since we defined the normal vector such that  $|\mathbf{n}_{pf}| = |e_f|$ . Similarly to (9), it can be shown that this is a second order approximation of the face integral. See, e.g., [5, 3].

We denote the volumetric flux as

$$a_{pf} \equiv \mathbf{u}(\mathbf{x}_{pf}) \cdot \mathbf{n}_{pf}.$$

Thus, we end up with the approximation of the divergence term

$$\int_{\Omega_p} \nabla \cdot (\phi \mathbf{u}) \approx \sum_{f \in \mathcal{F}_p} \phi(\mathbf{x}_{pf}, t) a_{pf}.$$

In order to achieve second order in time, we approximate the time integral of the fluxes using the midpoint rule

$$\int_{t^{n-1}}^{t^n} \sum_{f \in \mathcal{F}_p} \phi(\mathbf{x}_{pf}, t) a_{pf} \approx \Delta t \sum_{f \in \mathcal{F}_p} \phi(\mathbf{x}_{pf}, t^{n-1/2}) a_{pf} = \Delta t \sum_{f \in \mathcal{F}_p} \phi_{pf}^{n-1/2} a_{pf},$$

where  $\Delta t = t^n - t^{n-1}$ , is the uniform time step,  $t^{n-1/2} = \frac{t^n + t^{n-1}}{2}$ , and we use the notation  $\phi(\mathbf{x}_{pf}, t^n) = \phi_{pf}^n$ .

Thus, the second order in space and time discretization of (7) reads as

$$|\Omega_p|(\phi_p^n - \phi_p^{n-1}) + \Delta t \sum_{f \in \mathcal{F}_p} \phi_{pf}^{n-1/2} a_{pf} = 0. \quad (10)$$

Next we express the value at the face center at half-time step as the average value of old and new values. We take the average of reconstructed face values from neighboring cells, the upwind value is treated implicitly, and the downwind explicitly. In particular, for the outflow case  $a_{pf} \geq 0$

$$\phi_{pf}^{n-1/2} = \frac{1}{2}(\phi_{pf}^n + \phi_{qf}^{n-1}),$$

where

$$\phi_{pf}^n = \phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf}, \quad \text{and} \quad \phi_{qf}^{n-1} = \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}.$$

This mimics the approximation of the face value of the IIOE method [4, 1, 2] in 1D:

$$\phi_{i+1/2}^{n-1/2} = \frac{1}{2}(\phi_i^n + \phi_{i+1}^{n-1}).$$

For the inflow case  $a_{pf} < 0$

$$\phi_{pf}^{n-1/2} = \frac{1}{2}(\phi_{pf}^{n-1} + \phi_{qf}^n),$$

where

$$\phi_{pf}^{n-1} = \phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf}, \quad \text{and} \quad \phi_{qf}^n = \phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf}. \quad (11)$$

We divide the set of faces to inflow and outflow parts

$$\begin{aligned}\mathcal{B}_p^- &= \{b \in \mathcal{B}_p \mid a_{pb} < 0\} & \text{and} & & \mathcal{B}_p^+ &= \mathcal{B}_p \setminus \mathcal{B}_p^- \\ \mathcal{F}_p^- &= \{f \in (\mathcal{F}_p \setminus \mathcal{B}_p) \mid a_{pf} < 0\} & \text{and} & & \mathcal{F}_p^+ &= (\mathcal{F}_p \setminus \mathcal{B}_p) \setminus \mathcal{F}_p^-.\end{aligned}$$

Thus, we can rewrite (10)

$$\begin{aligned}\frac{|\Omega_p|}{\Delta t}(\phi_p^n - \phi_p^{n-1}) &+ \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} + \sum_{f \in \mathcal{F}_p^+} \phi_{pf}^{n-1/2} a_{pf} \\ &+ \sum_{b \in \mathcal{B}_p^-} \phi_{pb}^{n-1/2} a_{pb} + \sum_{b \in \mathcal{B}_p^+} \phi_{pb}^{n-1/2} a_{pb} = 0,\end{aligned}\tag{12}$$

where the face values depend on the sign of flux  $a_{pf}$

$$\phi_{pf}^{n-1/2} = \begin{cases} \frac{1}{2}(\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}) & \text{if } a_{pf} \geq 0 \\ \frac{1}{2}(\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf} + \phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf}) & \text{if } a_{pf} < 0, \end{cases}\tag{13}$$

$$\phi_{pb}^{n-1/2} = \begin{cases} \frac{1}{2}(\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_e(\mathbf{x}_{pb}, t^{n-1})) & \text{if } a_{pb} \geq 0 \\ \frac{1}{2}(\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf} + \phi_e(\mathbf{x}_{pb}, t^n)) & \text{if } a_{pb} < 0. \end{cases}\tag{14}$$

Substituting (13), (14) to (10) yields a linear system of equations

$$\begin{aligned}\frac{|\Omega_p|}{\Delta t}(\phi_p^n - \phi_p^{n-1}) &+ \sum_{f \in \mathcal{F}_p^-} \frac{1}{2}(\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf} + \phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf}) a_{pf} \\ &+ \sum_{f \in \mathcal{F}_p^+} \frac{1}{2}(\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}) a_{pf} \\ &+ \sum_{b \in \mathcal{B}_p^-} \frac{1}{2}(\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf} + \phi_e(\mathbf{x}_{pb}, t^n)) a_{pb} \\ &+ \sum_{b \in \mathcal{B}_p^+} \frac{1}{2}(\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_e(\mathbf{x}_{pb}, t^{n-1})) a_{pb} = 0.\end{aligned}\tag{15}$$

As described in [1], because of practical considerations, the simplest overlapping domain decomposition is required under the 1-ring face neighborhood structure in order to simplify the parallel computations. From a cell it is possible to access the values from neighboring cells. However, because of the unknown gradient terms of the neighbors  $\nabla \phi_q^n$ , we would need values from a 2-ring neighborhood. Another complication is that for the gradient term  $\nabla \phi_p^n$  it is quite cumbersome to extract the coefficients needed for the neighbors. So, because of the above practical considerations, we employ a deferred correction method [5, 1] to solve the system.

For the case  $a_{pf} \geq 0$  we can rewrite the approximation of the face values as

$$\begin{aligned}\phi_{pf}^{n-1/2} &= \frac{1}{2}(\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}) \\ &= \phi_p^n + \frac{1}{2}(-\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}),\end{aligned}\tag{16}$$

which is now written as the upwind cell center value plus a correction term. To simplify the notation, we denote the correction term as

$$\mathcal{D}_{pf}^+ \phi \equiv \frac{1}{2}(-\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}),$$

the plus sign indicating that  $a_{pf} \geq 0$  is non-negative. To ensure boundedness, we apply a limiter  $\Psi_f$  to the correction term

$$\phi_{pf}^{n-1/2} = \phi_p^n + \Psi_f \mathcal{D}_{pf}^+ \phi\tag{17}$$

For the case  $a_{pf} < 0$  we can rewrite the approximation of the face values as

$$\begin{aligned}\phi_{pf}^{n-1/2} &= \frac{1}{2}(\phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf} + \phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf}) \\ &= \phi_q^n + \frac{1}{2}(-\phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf} + \phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf}),\end{aligned}\tag{18}$$

which is now written, as in the outflow case, as the upwind cell center value plus a correction term. To simplify the notation, we denote the correction term as

$$\mathcal{D}_{pf}^- \phi \equiv \frac{1}{2}(-\phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf} + \phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf}),$$

the minus sign indicating that  $a_{pf} < 0$  is negative.

We do similarly at boundary faces. For  $a_{pb} \geq 0$  we have

$$\begin{aligned}\phi_{pb}^{n-1/2} &= \frac{1}{2}(\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pb} + \phi_e(\mathbf{x}_{pb}, t^{n-1})) \\ &= \phi_p^n + \frac{1}{2}(-\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pb} + \phi_e(\mathbf{x}_{pb}, t^{n-1})) \\ &= \phi_p^n + \mathcal{D}_{pb}^+ \phi,\end{aligned}$$

where

$$\mathcal{D}_{pb}^+ \phi = \frac{1}{2}(-\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pb} + \phi_e(\mathbf{x}_{pb}, t^{n-1})).$$

For  $a_{pb} < 0$  we have

$$\begin{aligned}\phi_{pb}^{n-1/2} &= \frac{1}{2}(\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pb} + \phi_e(\mathbf{x}_{pb}, t^n)) \\ &= \phi_e(\mathbf{x}_{pb}, t^n) + \frac{1}{2}(\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pb} - \phi_e(\mathbf{x}_{pb}, t^n)) \\ &= \phi_e(\mathbf{x}_{pb}, t^n) + \mathcal{D}_{pb}^- \phi,\end{aligned}$$

We end up with a linear system of equations

$$\begin{aligned}
& \frac{|\Omega_p|}{\Delta t} (\phi_p^n - \phi_p^{n-1}) + \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^n + \frac{\Psi_f}{2} (-\phi_q^n + \nabla \phi_q^n \cdot \mathbf{d}_{qf} + \phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pf}) \right) a_{pf} \\
& + \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^n + \frac{\Psi_f}{2} (-\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pf} + \phi_q^{n-1} + \nabla \phi_q^{n-1} \cdot \mathbf{d}_{qf}) \right) a_{pf} \\
& + \sum_{b \in \mathcal{B}_p^-} \left( \phi_e(\mathbf{x}_{pb}, t^n) + \frac{\Psi_b}{2} (\phi_p^{n-1} + \nabla \phi_p^{n-1} \cdot \mathbf{d}_{pb} - \phi_e(\mathbf{x}_{pb}, t^n)) \right) a_{pb} \\
& + \sum_{b \in \mathcal{B}_p^+} \left( \phi_p^n + \frac{\Psi_b}{2} (-\phi_p^n + \nabla \phi_p^n \cdot \mathbf{d}_{pb} + \phi_e(\mathbf{x}_{pb}, t^{n-1})) \right) a_{pb} = 0.
\end{aligned} \tag{19}$$

We solve the system iteratively by applying the deferred correction procedure [5] for similar reasons as in [1]. In order to keep the parallel computations simple, in each cell a 1-ring neighborhood structure should be considered only. However, to compute the gradient of a neighbor cell  $q$ , we would need access to cells outside the 1-ring neighborhood. Thus, the gradients of neighbors  $q$  should be computed using the latest available values from the previous iteration. For the case  $a_{pf} \geq 0$ , the face value in the  $k$ -th iteration becomes

$$\phi_{pf}^{n-1/2,k} = \phi_p^{n,k} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \tag{20}$$

$$\begin{aligned}
& \frac{|\Omega_p|}{\Delta t} (\phi_p^{n,k} - \phi_p^{n-1}) + \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k} + \Psi_{pf}^{-,k-1} \mathcal{D}_{pf}^- \phi^{k-1} \right) a_{pf} \\
& + \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k} + \Psi_{pf}^{+,k-1} \mathcal{D}_{pf}^+ \phi^{k-1} \right) a_{pf} \\
& + \sum_{b \in \mathcal{B}_p^-} \left( \phi_e(\mathbf{x}_{pb}, t^n) + \Psi_{pb}^- \mathcal{D}_{pb}^- \phi \right) a_{pb} \\
& + \sum_{b \in \mathcal{B}_p^+} \left( \phi_p^{n,k} + \Psi_{pb}^{+,k-1} \mathcal{D}_{pb}^+ \phi^{k-1} \right) a_{pb} = 0.
\end{aligned} \tag{21}$$

This yields a system with an M-matrix, which has favorable properties ([reference here](#)). The inverse of an M-matrix has positive components. We can think of the solution of the system as applying the inverse matrix  $\mathbf{M}^{-1}$  to the RHS.

We rearrange the terms

$$\begin{aligned}
& \frac{|\Omega_p|}{\Delta t} \phi_p^{n,k} + \sum_{f \in \mathcal{F}_p^-} \phi_q^{n,k} a_{pf} + \sum_{f \in \mathcal{F}_p^+} \phi_p^{n,k} a_{pf} + \sum_{b \in \mathcal{B}_p^+} \phi_p^{n,k} a_{pb} = \\
& \frac{|\Omega_p|}{\Delta t} \phi_p^{n-1} - \sum_{f \in \mathcal{F}_p^-} \Psi_{pf}^{-,k-1} \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} - \sum_{f \in \mathcal{F}_p^+} \Psi_{pf}^{+,k-1} \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf} \\
& - \sum_{b \in \mathcal{B}_p^-} (\phi_e(\mathbf{x}_{pb}, t^n) + \Psi_{pb}^- \mathcal{D}_{pb}^- \phi) a_{pb} - \sum_{b \in \mathcal{B}_p^+} \Psi_{pb}^{+,k-1} \mathcal{D}_{pb}^+ \phi^{k-1} a_{pb}.
\end{aligned} \tag{22}$$

For internal cells, we require for the right-hand side

$$\phi_{p,vnbd}^{min} \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \Psi_{pf}^{-,k-1} \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \Psi_{pf}^{+,k-1} \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf} \leq \phi_{p,vnbd}^{max}.$$

Let us define

$$\phi_{est,p} = \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf}. \tag{23}$$

First we check in each cell if  $\phi_{est,p} > \phi_{p,vnbd}^{max}$ , then we define

$$\phi_{target,p} = \phi_{p,vnbd}^{max},$$

else if  $\phi_{est,p} < \phi_{p,vnbd}^{min}$ , then

$$\phi_{target,p} = \phi_{p,vnbd}^{min},$$

otherwise

$$\phi_{target,p} = \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf}.$$

We can write this in a compact form as

$$\phi_{target,p}^n = \min(\phi_{p,vnbd}^{max}, \max(\phi_{p,vnbd}^{min}, \phi_{est,p}^n)). \tag{24}$$

We assume, that

$$\phi_{target,p}^n = \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \Psi_{pf}^{-,k-1} \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \Psi_{pf}^{+,k-1} \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf}, \tag{25}$$

$$\sum_{f \in \mathcal{F}_p^-} \Psi_{pf}^{-,k-1} \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} + \sum_{f \in \mathcal{F}_p^+} \Psi_{pf}^{+,k-1} \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf} = \frac{\phi_p^{n-1} - \phi_{target,p}^n}{\lambda_p}. \tag{26}$$

Furthermore, we want to minimize

$$\sum_{f \in \mathcal{F}_p^-} \left( \Psi_{pf}^{-,k-1} - 1 \right)^2 \left( \mathcal{D}_{pf}^- \phi^{k-1} a_{pf} \right)^2 + \sum_{f \in \mathcal{F}_p^+} \left( \Psi_{pf}^{+,k-1} - 1 \right)^2 \left( \mathcal{D}_{pf}^+ \phi^{k-1} a_{pf} \right)^2 \tag{27}$$



$$\phi_{p,vnbd}^{min} \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_{pf}^{n-1/2} a_{pf} \leq \phi_{p,vnbd}^{max}$$

For  $a_{ph} \geq 0$

$$\phi_{p,vnbd}^{min} \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \lambda_p \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} - \lambda_p \phi_{ph}^{n-1/2} a_{ph} \leq \phi_{p,vnbd}^{max}$$

$$\phi_{p,vnbd}^{min} - \phi_p^{n-1} + \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} + \lambda_p \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \leq -\lambda_p \phi_{ph}^{n-1/2} a_{ph}$$

$$\phi_{p,vnbd}^{max} - \phi_p^{n-1} + \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} + \lambda_p \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \geq -\lambda_p \phi_{ph}^{n-1/2} a_{ph}$$

$$\frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \geq \phi_{ph}^{n-1/2} a_{ph}$$

$$\frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \leq \phi_{ph}^{n-1/2} a_{ph}$$

## 0.1 Version 1

Assume

$$0 \leq \phi_{p,vnbd}^{min} \leq \phi_{pf}^{n-1/2} \leq \phi_{p,vnbd}^{max} \leq 1.$$

Then

$$\phi_{p,vnbd}^{min} \sum_{f \in \mathcal{F}_p^+} a_{pf} \leq \sum_{f \in \mathcal{F}_p^+} \phi_{pf}^{n-1/2} a_{pf} \leq \phi_{p,vnbd}^{max} \sum_{f \in \mathcal{F}_p^+} a_{pf},$$

$$\phi_{p,vnbd}^{min} \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf} \leq \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \leq \phi_{p,vnbd}^{max} \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf}$$

and

$$\phi_{p,vnbd}^{min} \sum_{f \in \mathcal{F}_p^-} a_{pf} \geq \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} \geq \phi_{p,vnbd}^{max} \sum_{f \in \mathcal{F}_p^-} a_{pf}.$$

Thus,

$$\begin{aligned} & \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \geq \\ & \geq \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \phi_{p,vnbd}^{min} \sum_{f \in \mathcal{F}_p^-} a_{pf} - \phi_{p,vnbd}^{max} \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf} \geq \phi_{ph}^{n-1/2} a_{ph}. \end{aligned}$$

Similarly,

$$\begin{aligned}
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \leq \\
& \leq \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \phi_{p,vnbd}^{max} \sum_{f \in \mathcal{F}_p^-} a_{pf} - \phi_{p,vnbd}^{min} \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf} \leq \phi_{ph}^{n-1/2} a_{ph}. \\
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \phi_{p,vnbd}^{min} \sum_{f \in \mathcal{F}_p^-} a_{pf} - \phi_{p,vnbd}^{max} \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf} \geq \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \phi_{p,vnbd}^{max} \sum_{f \in \mathcal{F}_p^-} a_{pf} - \phi_{p,vnbd}^{min} \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf} \\
& 0 \leq \frac{\phi_{p,vnbd}^{max} - \phi_{p,vnbd}^{min}}{\lambda_p} + (\phi_{p,vnbd}^{max} - \phi_{p,vnbd}^{max}) \sum_{f \in \mathcal{F}_p^-} a_{pf} - (\phi_{p,vnbd}^{max} - \phi_{p,vnbd}^{max}) \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} a_{pf}
\end{aligned}$$

## 0.2 Version 2

Let  $\min(\phi_p^{n-1}, \phi_q^{n-1}) = \min_{pq}^{n-1}$  and  $\max(\phi_p^{n-1}, \phi_q^{n-1}) = \max_{pq}^{n-1}$ . Assume

$$0 \leq \min_{pq}^{n-1} \leq \phi_{pf}^{n-1/2} \leq \max_{pq}^{n-1} \leq 1.$$

Then

$$\sum_{f \in \mathcal{F}_p^+} \min_{pq}^{n-1} a_{pf} \leq \sum_{f \in \mathcal{F}_p^+} \phi_{pf}^{n-1/2} a_{pf} \leq \sum_{f \in \mathcal{F}_p^+} \max_{pq}^{n-1} a_{pf},$$

and

$$\begin{aligned}
& \sum_{f \in \mathcal{F}_p^-} \max_{pq}^{n-1} a_{pf} \leq \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} \leq \sum_{f \in \mathcal{F}_p^-} \min_{pq}^{n-1} a_{pf}. \\
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \geq \\
& \geq \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \min_{pq}^{n-1} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \max_{pq}^{n-1} a_{pf} \geq \phi_{ph}^{n-1/2} a_{ph}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \phi_{pf}^{n-1/2} a_{pf} \leq \\
& \leq \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \max_{pq}^{n-1} a_{pf} - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \min_{pq}^{n-1} a_{pf} \leq \phi_{ph}^{n-1/2} a_{ph}.
\end{aligned}$$

### 0.3 Version 3

$$\begin{aligned}
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} \\
& - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \geq \left( \phi_p^{n,k-1} + \Psi_h^k \mathcal{D}_{ph}^{+,k-1} \phi \right) a_{ph}, \\
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{max}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} \\
& - \sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \leq \left( \phi_p^{n,k-1} + \Psi_h^k \mathcal{D}_{ph}^{+,k-1} \phi \right) a_{ph}.
\end{aligned}$$

We want to maximize  $\Psi$  at each face.

$$\sum_{\substack{f \in \mathcal{F}_p^+, \\ f \neq h}} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} = \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} - \left( \phi_p^{n,k-1} + \Psi_h^{k-1} \mathcal{D}_{ph}^{+,k-1} \phi \right) a_{ph}.$$

Substituting we get

$$\begin{aligned}
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} \\
& - \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \\
& + \left( \phi_p^{n,k-1} + \Psi_h^{k-1} \mathcal{D}_{ph}^{+,k-1} \phi \right) a_{ph} \geq \left( \phi_p^{n,k-1} + \Psi_h^k \mathcal{D}_{ph}^{+,k-1} \phi \right) a_{ph}, \\
& \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} \\
& - \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \\
& + \Psi_h^{k-1} \mathcal{D}_{ph}^{+,k-1} \phi a_{ph} \geq \Psi_h^k \mathcal{D}_{ph}^{+,k-1} \phi a_{ph}, \\
& \Psi_h^k = \max \left( \frac{\phi_p^{n-1} - \phi_{p,vnbd}^{min}}{\lambda_p \mathcal{D}_{ph}^{+,k-1} \phi a_{ph}} - \frac{\sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf}}{\mathcal{D}_{ph}^{+,k-1} \phi a_{ph}} \right. \\
& \quad \left. - \frac{\sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf}}{\mathcal{D}_{ph}^{+,k-1} \phi a_{ph}} + \Psi_h^{k-1}, \right)
\end{aligned}$$

## 0.4 Version 4

First we make an estimate for the new cell value, using the maximal possible value for the limiter at outflow faces.

$$\phi_{est}^n = \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_{max} \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf}$$

Then we compute a bounded value

$$\phi_{bounded}^n = \min \left( \phi_{p,vnbd}^{max}, \max \left( \phi_{p,vnbd}^{min}, \phi_{est}^n \right) \right) \quad (28)$$

By assuming

$$\phi_{bounded}^n = \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \quad (29)$$

we can compute the sum

$$\begin{aligned} S_p &= \sum_{f \in \mathcal{F}_p^+} \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} \\ &= \frac{\phi_p^{n-1} - \phi_{bounded}^n}{\lambda_p} - \sum_{f \in \mathcal{F}_p^-} \left( \phi_q^{n,k-1} + \Psi_f^{k-1} \mathcal{D}_{pf}^{-,k-1} \phi \right) a_{pf} - \sum_{f \in \mathcal{F}_p^+} \phi_p^{n,k-1} a_{pf}. \end{aligned} \quad (30)$$

We want to be as close to the second order scheme as possible. Thus, we want to minimize

$$\begin{aligned} &\left( \sum_{f \in \mathcal{F}_p^+} \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} - \Psi_{max} \sum_{f \in \mathcal{F}_p^+} \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} \right)^2 = \\ &\left( \sum_{f \in \mathcal{F}_p^+} (\Psi_f^k - \Psi_{max}) \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} \right)^2. \end{aligned}$$

This, however, is not good, because we already fixed the sum to  $S_p$ !!!

Be careful, that

$$\left( \sum_{f \in \mathcal{F}_p^+} (\Psi_f^k - \Psi_{max}) \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} \right)^2 \neq \sum_{f \in \mathcal{F}_p^+} (\Psi_f^k - \Psi_{max})^2 \left( \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} \right)^2$$

!!!!!!

Minimizing

$$\sum_{f \in \mathcal{F}_p^+} (\Psi_f^k - \Psi_{max})^2 \left( \mathcal{D}_{pf}^{+,k-1} \phi a_{pf} \right)^2 \quad (31)$$

seems reasonable.

$$\begin{aligned} \mathcal{L}(\dots, \Psi_f^k, \dots, \lambda) = & \sum_{f \in \mathcal{F}_p^+} (\Psi_f^k - \Psi_{max})^2 \left( \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right)^2 \\ & - \lambda \left( \sum_{f \in \mathcal{F}_p^+} \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} - S_p \right) \end{aligned} \quad (32)$$

$$\frac{\partial \mathcal{L}}{\partial \Psi_f^k} = 0$$

$$\begin{aligned} 2 (\Psi_f^k - \Psi_{max}) \left( \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right)^2 - \lambda \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} &= 0 \\ \Psi_f^k &= \Psi_{max} + \frac{\lambda}{2 \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf}} \end{aligned}$$

We know that

$$\begin{aligned} S_p &= \sum_{f \in \mathcal{F}_p^+} \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \\ &= \sum_{f \in \mathcal{F}_p^+} \left( \Psi_{max} + \frac{\lambda}{2 \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf}} \right) \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \\ &= \sum_{f \in \mathcal{F}_p^+} \Psi_{max} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} + \frac{\lambda}{2} n_{pos}, \end{aligned} \quad (33)$$

where  $n_{pos} = |\mathcal{F}_p^+|$ . We can compute  $\lambda$

$$\lambda = \frac{2}{n_{pos}} \left( S_p - \Psi_{max} \sum_{f \in \mathcal{F}_p^+} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right) \quad (34)$$

$$\Psi_f^k = \max \left( \min \left( \Psi_{max} + \frac{\lambda}{2 \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf}}, \Psi_{max} \right), 0 \right) \quad (35)$$

Another choice would be to minimize

$$\begin{aligned} & \sum_{f \in \mathcal{F}_p^+} (\Psi_f^k - \Psi_{max})^2 \\ & \frac{\partial \mathcal{L}}{\partial \Psi_f^k} = 0 \\ 2 (\Psi_f^k - \Psi_{max}) - \lambda \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} &= 0 \\ \Psi_f^k &= \Psi_{max} + \frac{\lambda}{2} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \end{aligned}$$

We know that

$$\begin{aligned}
S_p &= \sum_{f \in \mathcal{F}_p^+} \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \\
&= \sum_{f \in \mathcal{F}_p^+} \left( \Psi_{max} + \frac{\lambda}{2} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right) \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \\
&= \sum_{f \in \mathcal{F}_p^+} \Psi_{max} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} + \frac{\lambda}{2} \sum_{f \in \mathcal{F}_p^+} \left( \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right)^2,
\end{aligned} \tag{36}$$

$$\lambda = \frac{2}{\sum_{f \in \mathcal{F}_p^+} \left( \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right)^2} \left( S_p - \Psi_{max} \sum_{f \in \mathcal{F}_p^+} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf} \right) \tag{37}$$

$$\Psi_f^k = \max \left( \min \left( \Psi_{max} + \frac{\lambda}{2} \mathcal{D}_{pf}^{+,k-1} \phi \ a_{pf}, \Psi_{max} \right), 0 \right) \tag{38}$$

## 0.5 Version 5

$$\begin{aligned}
\phi_{p,vnbd}^{min} \sum_{f \in \mathcal{F}_p^-} a_{pf} &\geq \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} \geq \phi_{p,vnbd}^{max} \sum_{f \in \mathcal{F}_p^-} a_{pf}. \\
\phi_{p,vnbd}^{min} &\leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_{pf}^{n-1/2} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \leq \phi_{p,vnbd}^{max} \\
\phi_{p,vnbd}^{min} &\leq \phi_p^{n-1} - \lambda_p \phi_{p,vnbd}^{min} \sum_{f \in \mathcal{F}_p^-} a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \\
\phi_{p,vnbd}^{min} \left( 1 + \lambda_p \sum_{f \in \mathcal{F}_p^-} a_{pf} \right) &\leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \left( \phi_p^{n,k-1} + \Psi_f^k \mathcal{D}_{pf}^{+,k-1} \phi \right) a_{pf} \\
\phi_{p,vnbd}^{min} \left( 1 + \lambda_p \sum_{f \in \mathcal{F}_p^-} a_{pf} \right) &\leq \phi_{p,vnbd}^{max} \left( 1 + \lambda_p \sum_{f \in \mathcal{F}_p^-} a_{pf} \right)
\end{aligned}$$

Only holds if

$$1 + \lambda_p \sum_{f \in \mathcal{F}_p^-} a_{pf} > 0.$$

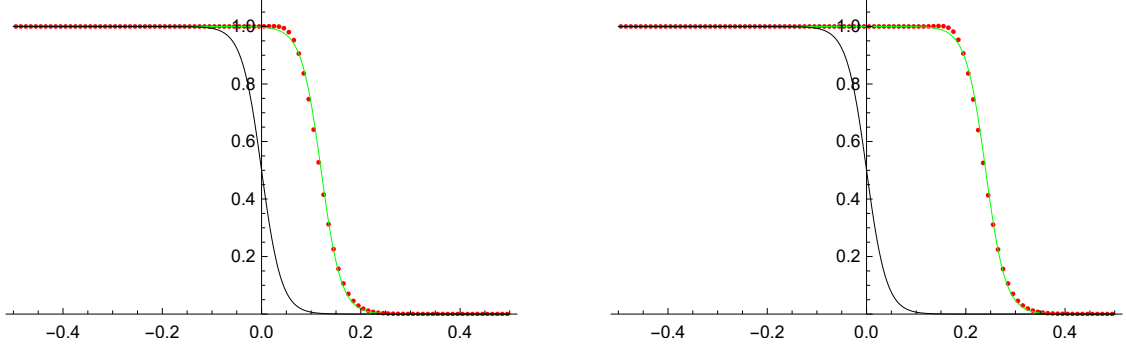


Figure 1: Comparing the

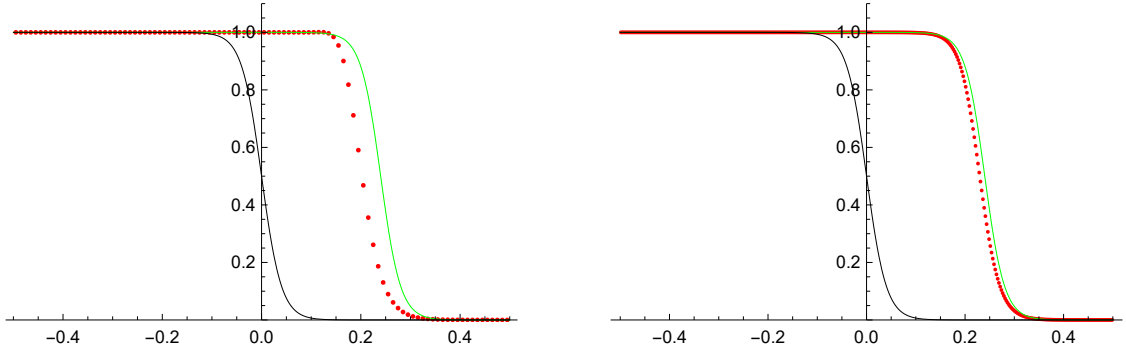


Figure 2: Comparing the solution is getting closer to the exact speed.

$$\begin{aligned}
& \frac{|\Omega_p|}{\Delta t} (\phi_p^n - \phi_p^{n-1}) + \sum_{f \in \mathcal{F}_p^-} (\phi_q^n + \Psi_{pf}^- \mathcal{D}_{pf}^- \phi) a_{pf} \\
& + \sum_{f \in \mathcal{F}_p^+} (\phi_p^n + \Psi_{pf}^+ \mathcal{D}_{pf}^+ \phi) a_{pf} \\
& + \sum_{b \in \mathcal{B}_p^-} (\phi_e(\mathbf{x}_{pb}, t^n) + \Psi_{pb}^- \mathcal{D}_{pb}^- \phi) a_{pb} \\
& + \sum_{b \in \mathcal{B}_p^+} (\phi_p^n + \Psi_{pb}^+ \mathcal{D}_{pb}^+ \phi) a_{pb} = 0.
\end{aligned} \tag{39}$$

For an internal cell

$$\begin{aligned}
& \frac{|\Omega_p|}{\Delta t} (\phi_p^n - \phi_p^{n-1}) + \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} + \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} \\
& + \sum_{f \in \mathcal{F}_p^-} \Psi_f \mathcal{D}_{pf}^- \phi a_{pf} + \sum_{f \in \mathcal{F}_p^+} \Psi_f \mathcal{D}_{pf}^+ \phi a_{pf} = 0.
\end{aligned} \tag{40}$$

We split the antidiffusive fluxes to positive and negative values. Then we can write the sum of antidiffusive fluxes as

$$\begin{aligned}
& \sum_{f \in \mathcal{F}_p^-} \Psi_f \mathcal{D}_{pf}^- \phi a_{pf} + \sum_{f \in \mathcal{F}_p^+} \Psi_f \mathcal{D}_{pf}^+ \phi a_{pf} = \\
& \sum_{f \in \mathcal{F}_p^+} \Psi_f \max(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^+} \Psi_f \min(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) \\
& + \sum_{f \in \mathcal{F}_p^-} \Psi_f \max(\mathcal{D}_{pf}^- \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \Psi_f \min(\mathcal{D}_{pf}^- \phi a_{pf}, 0).
\end{aligned} \tag{41}$$

Since the limiter is positive  $\Psi_f \geq 0$ , we can write

$$\begin{aligned}
& \sum_{f \in \mathcal{F}_p^+} \Psi_f \max(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \Psi_f \max(\mathcal{D}_{pf}^- \phi a_{pf}, 0) = \\
& \Psi_p^M \left( \sum_{f \in \mathcal{F}_p^+} \max(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \max(\mathcal{D}_{pf}^- \phi a_{pf}, 0) \right),
\end{aligned} \tag{42}$$

where  $\Psi_p^M \geq 0$ . Similarly

$$\begin{aligned}
& \sum_{f \in \mathcal{F}_p^+} \Psi_f \min(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \Psi_f \min(\mathcal{D}_{pf}^- \phi a_{pf}, 0) = \\
& \Psi_p^N \left( \sum_{f \in \mathcal{F}_p^+} \min(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \min(\mathcal{D}_{pf}^- \phi a_{pf}, 0) \right),
\end{aligned} \tag{43}$$

where  $\Psi_p^N \geq 0$ . We denote

$$\begin{aligned}
M_p &= \sum_{f \in \mathcal{F}_p^+} \max(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \max(\mathcal{D}_{pf}^- \phi a_{pf}, 0), \\
N_p &= \sum_{f \in \mathcal{F}_p^+} \min(\mathcal{D}_{pf}^+ \phi a_{pf}, 0) + \sum_{f \in \mathcal{F}_p^-} \min(\mathcal{D}_{pf}^- \phi a_{pf}, 0)
\end{aligned} \tag{44}$$

we can equivalently write

$$\frac{|\Omega_p|}{\Delta t} (\phi_p^n - \phi_p^{n-1}) + \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} + \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} + \Psi_p^M M_p + \Psi_p^N N_p = 0. \tag{45}$$

After finding the single value  $\Psi_p^M$ , we should find a way how do we distribute it among the faces to get the most accurate solution.

$$\phi_p^n - \phi_p^{n-1} + \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} + \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} + \lambda_p \Psi_p^M M_p + \lambda_p \Psi_p^N N_p = 0. \tag{46}$$



We want

$$\min \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \lambda_p \Psi_p^M M_p - \lambda_p \Psi_p^N N_p \leq \max$$

Since  $N_p \leq \Psi_p^N N_p \leq 0$  we can write the sufficient conditions

$$\phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \lambda_p \Psi_p^M M_p - \lambda_p N_p \leq \max$$

$$\min \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \lambda_p \Psi_p^M M_p.$$

Thus

$$\phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \lambda_p N_p - \max \leq \lambda_p \Psi_p^M M_p$$

and

$$\lambda_p \Psi_p^M M_p \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \min.$$

It is possible to bound  $\Psi_p^M$  if

$$\begin{aligned} \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \lambda_p N_p - \max &\leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \min \\ &- \lambda_p N_p \leq \max - \min \end{aligned}$$

We want

$$min \leq \phi_p^{n-1} - \lambda_p \sum_{f \in \mathcal{F}_p^-} \phi_q^n a_{pf} - \lambda_p \sum_{f \in \mathcal{F}_p^+} \phi_p^n a_{pf} - \alpha_p (\lambda_p M_p + \lambda_p N_p) \leq max$$

end if

# Bibliography

- [1] Jooyoung Hahn, Karol Mikula, Peter Frolkovič, Matej Medl’a, and Branislav Basara. “Iterative inflow-implicit outflow-explicit finite volume scheme for level-set equations on polyhedron meshes”. In: *Computers & Mathematics with Applications* 77.6 (2019), pp. 1639–1654. URL: [https://www.math.sk/mikula/hmfmb\\_cma\\_2018.pdf](https://www.math.sk/mikula/hmfmb_cma_2018.pdf).
- [2] Gergő Ibolya and Karol Mikula. “Numerical solution of the 1d viscous Burgers and traffic flow equations by the inflow-implicit/outflow-explicit finite volume method”. In: *Proceedings of ALGORITMY*. 2020, pp. 191–200. URL: [https://www.math.sk/mikula/ibolya\\_alg2020.pdf](https://www.math.sk/mikula/ibolya_alg2020.pdf).
- [3] Hrvoje Jasak. “Error analysis and estimation for the finite volume method with applications to fluid flows.” In: (1996).
- [4] Karol Mikula, Mario Ohlberger, and Jozef Urbán. “Inflow-implicit/outflow-explicit finite volume methods for solving advection equations”. English. In: *Applied Numerical Mathematics* 85 (2014), pp. 16–37. ISSN: 0168-9274. URL: [https://www.math.sk/mikula/mou\\_apnum.pdf](https://www.math.sk/mikula/mou_apnum.pdf).
- [5] Fadl Moukalled, Luca Mangani, Marwan Darwish, F Moukalled, L Mangani, and M Darwish. *The finite volume method*. Springer, 2016. URL: <http://www.gidropraktikum.narod.ru/Moukalled-et-al-FVM-OpenFOAM-Matlab.pdf>.