



MOX-Report No. 26/2022

**A filtering monotonization approach for DG
discretizations of hyperbolic problems**

Orlando, G.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it

<http://mox.polimi.it>

A filtering monotonization approach for DG discretizations of hyperbolic problems

Giuseppe Orlando

MOX, Dipartimento di Matematica, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
`giuseppe.orlando@polimi.it`

Keywords: Discontinuous Galerkin method, monotone schemes, conservation laws, strong stability preserving methods, filtering methods

AMS Subject Classification: 35A16, 35L65, 35L67, 65M50, 65M60

Abstract

We introduce a filtering technique for Discontinuous Galerkin approximations of hyperbolic problems. Following an approach already proposed for the Hamilton-Jacobi equations by other authors, we aim at reducing the spurious oscillations that arise in presence of discontinuities when high order spatial discretizations are employed. This goal is achieved using a filter function that keeps the high order scheme when the solution is regular and switches to a monotone low order approximation if it is not. The method has been implemented in the framework of the *deal.II* numerical library, whose mesh adaptation capabilities are also used to reduce the region in which the low order approximation is used. A number of numerical experiments demonstrate the potential of the proposed filtering technique.

1 Introduction

The Discontinuous Galerkin (DG) method has proven itself a very valuable tool for applications to computational fluid dynamics problems in a great variety of flow regimes, see e.g. the seminal contributions [3, 4, 8, 7, 6, 9, 10] as well as the reviews in [14, 17], among many others. For hyperbolic problems, however, spurious oscillations can arise around shocks and other discontinuities when high order spatial discretizations are used. Furthermore, in many applications, maintaining non negativity of the numerical solutions is essential to preserve their correct physical meaning. In order to address these well known issues, a number of monotization techniques have been proposed in the literature for DG methods. While a full survey of this topic goes beyond the scope of the paper, we review briefly here some of the most popular techniques.

In general, monotization techniques for DG methods have been inherited from finite difference and finite volume approaches. For example, starting with [8, 7, 9], slope limiting techniques have been employed, while other authors have investigated WENO methods [25, 26] and flux corrected transport methods [19, 22, 18]. Another approach is based on the identification of the regions where the discontinuities are located, in which the mesh is then refined and/or the order of the spatial discretization is lowered in order to exploit the monotonicity of (most) low order approximations. In recent years, the very successful MOOD approach has been proposed in [11, 12, 21, 30], which is also based on the identification of the regions of discontinuity and on the switch from a high order DG method to a monotonic first order finite volume method on a locally refined mesh built around the quadrature nodes used by the DG method.

The method proposed in this paper is inspired by the filtering approach outlined in [5, 24]. More specifically, a filter function is employed in such a way that, where the solution is regular, we keep the high order solution, whereas otherwise we switch to a low order method. While the proposed strategy is conceptually similar to that of the MOOD approach, the main novelty of the proposed method is that we do not rely on a regularity indicator and that a monotonic solution is retrieved (almost) automatically. For the spatial discretization, we use the DG approach implemented in the numerical library *deal.II* [2], which provides h -refinement capabilities that are exploited in order to reduce the size of the region where the low order approximation is applied.

The model problem is introduced in Section 2, along with the space and time discretizations that will be employed. The filtering mono-

tonization approach is introduced in Section 3 for a scalar hyperbolic PDE and extended in Section 4 to the inviscid Euler equations. Numerical results validating the proposed approach are presented in Section 5, while some conclusions and perspectives for future work are presented in Section 6.

2 Model problem and discretization

We consider as model problem the nonlinear conservation law

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{F}(u) = 0, \quad (1)$$

where $\mathbf{F}(u)$ denotes a d -dimensional vector field that depends on the unknown $u = u(\mathbf{x}, t)$, and $\mathbf{x} \in \mathbb{R}^d$, generally in a non linear way. Simple examples are the linear advection equation and the Burgers equation. We consider a decomposition of the domain Ω into a family of quadrilaterals \mathcal{T}_h , where each element is denoted by K . The skeleton \mathcal{E} denotes the set of all element faces and $\mathcal{E} = \mathcal{E}^I \cup \mathcal{E}^B$, where \mathcal{E}^I is the subset of interior faces and \mathcal{E}^B is the subset of boundary faces. Suitable jump and average operators can then be defined as customary for Discontinuous Galerkin discretizations. A face $\Gamma \in \mathcal{E}^I$ shares two elements that we denote by K^+ with outward unit normal \mathbf{n}^+ and K^- with outward unit normal \mathbf{n}^- , whereas for a face $\Gamma \in \mathcal{E}^B$ we denote by \mathbf{n} the outward unit normal. For a scalar function φ the jump is defined as

$$[[\varphi]] = \varphi^+ \mathbf{n}^+ + \varphi^- \mathbf{n}^- \quad \text{if } \Gamma \in \mathcal{E}^I \quad [[\varphi]] = \varphi \mathbf{n} \quad \text{if } \Gamma \in \mathcal{E}^B.$$

The average is defined as

$$\{\{\varphi\}\} = \frac{1}{2} (\varphi^+ + \varphi^-) \quad \text{if } \Gamma \in \mathcal{E}^I \quad \{\{\varphi\}\} = \varphi \quad \text{if } \Gamma \in \mathcal{E}^B.$$

Similar definitions apply for a vector function $\boldsymbol{\varphi}$:

$$\begin{aligned} [[\boldsymbol{\varphi}]] &= \boldsymbol{\varphi}^+ \cdot \mathbf{n}^+ + \boldsymbol{\varphi}^- \cdot \mathbf{n}^- \quad \text{if } \Gamma \in \mathcal{E}^I & [[\boldsymbol{\varphi}]] &= \boldsymbol{\varphi} \cdot \mathbf{n} \quad \text{if } \Gamma \in \mathcal{E}^B \\ \{\{\boldsymbol{\varphi}\}\} &= \frac{1}{2} (\boldsymbol{\varphi}^+ + \boldsymbol{\varphi}^-) \quad \text{if } \Gamma \in \mathcal{E}^I & \{\{\boldsymbol{\varphi}\}\} &= \boldsymbol{\varphi} \quad \text{if } \Gamma \in \mathcal{E}^B. \end{aligned}$$

We also introduce the following finite element spaces

$$Q_k = \{v \in L^2(\Omega) : v|_K \in \mathbb{Q}_k \quad \forall K \in \mathcal{T}_h\},$$

where \mathbb{Q}_k is the space of polynomials of degree k in each coordinate direction.

Concerning the time discretization, we will consider here only the well known TVD Runge Kutta methods described in [15, 16]. These are high order time discretization schemes that preserve the strong stability properties of first order explicit Euler time stepping and are known as Strong Stability Preserving (SSP) methods. For the convenience of the reader, we briefly recall here the second order and the third order optimal SSP Runge-Kutta methods derived in [15] for an ordinary differential equation $y' = \mathcal{N}(y)$. The second order scheme reads as follows:

$$v^{(1)} = v^n + \Delta t \mathcal{N}(v^n) \quad (2)$$

$$v^{n+1} = \frac{1}{2}v^n + \frac{1}{2}v^{(1)} + \frac{1}{2}\Delta t \mathcal{N}(v^{(1)}), \quad (3)$$

where $v^n \approx y(t^n)$ and Δt denotes the time step. The third order method is given instead by:

$$v^{(1)} = v^n + \Delta t \mathcal{N}(v^n) \quad (4)$$

$$v^{(2)} = \frac{3}{4}v^n + \frac{1}{4}v^{(1)} + \frac{1}{4}\Delta t \mathcal{N}(v^{(1)}) \quad (5)$$

$$v^{n+1} = \frac{1}{3}v^n + \frac{2}{3}v^{(2)} + \frac{2}{3}\Delta t \mathcal{N}(v^{(2)}). \quad (6)$$

Each stage of the TVD method can be represented as

$$\mathbf{u} = \mathbf{S}(\mathbf{v}), \quad (7)$$

where \mathbf{u}, \mathbf{v} denote the new and old values, respectively, of the vector containing the discrete degrees of freedom which identify the spatial approximation to the solutions of (1). \mathbf{S} denotes formally the solution operator associated to a specific time and space discretization. The transition from \mathbf{v} to \mathbf{u} can be interpreted as an advancement in time of $\alpha \Delta t$ time units, where α depends on the details of the TVD method and on the specific stage considered. We will denote by \mathbf{S}^M the discrete operator associated to the monotonic, low order spatial discretization and by \mathbf{S}^H that associated to a high order, not monotonic spatial discretization.

3 Outline of filtering monotonization approach

We will now introduce the application of the filtering approach proposed in [5, 24] in the above outlined context. First of all, a filter function F must be introduced. This can be defined in several ways, for example

$$F_1(x) = x\mathbf{1}_{|x|\leq 1}, \quad F_2(x) = \text{sign}(x) \max(1 - ||x| - 1|, 0). \quad (8)$$

In the simplest possible filtering approach, the filtered version of \mathbf{u} can be defined as

$$\mathbf{u}^F = \mathbf{S}^M(\mathbf{v}) + \varepsilon\alpha\Delta t F \left(\frac{\mathbf{S}^H(\mathbf{v}) - \mathbf{S}^M(\mathbf{v})}{\varepsilon\alpha\Delta t} \right), \quad (9)$$

where the low order solution \mathbf{S}^M is computed on the nodes of the high order solution \mathbf{S}^H and ε is a suitable parameter, depending on the time and space discretization parameters, such that

$$\lim_{(\Delta t, h) \rightarrow 0} \varepsilon(\Delta t, h) = 0,$$

with $h = \max\{\text{diam}(K) | K \in \mathcal{T}_h\}$. More details about the choice of ε will be given in Section 5. Notice that the filter function is applied componentwise. In this way, as discussed in [5], the high order method is only applied to the components i for which

$$\frac{|\mathbf{S}^H(\mathbf{v})_i - \mathbf{S}^M(\mathbf{v})_i|}{\varepsilon\alpha\Delta t} \leq 1, \quad i = 1, \dots, \dim(Q_h).$$

As explained in [5], ε has to be chosen in such a way that

$$\varepsilon \geq c_0 h,$$

where c_0 is a sufficiently large constant. As we will see in Section 5, the aforementioned approach is very dissipative and, unless a very large value of c_0 is adopted, it yields solutions that essentially coincide with the low order one. Therefore, we propose the alternative filtering strategy

$$\begin{aligned} \mathbf{u}_i^F &= \mathbf{S}^M(\mathbf{v})_i \\ &+ \beta \mathbf{S}^M(\mathbf{v})_i F \left(\frac{\mathbf{S}^H(\mathbf{v})_i - \mathbf{S}^M(\mathbf{v})_i}{\beta \mathbf{S}^M(\mathbf{v})_i} \right), \quad i = 1, \dots, \dim(Q_h), \end{aligned} \quad (10)$$

where $\beta > 0$ is a suitable parameter that represents a tolerance for the “componentwise relative difference” $\frac{\mathbf{S}^H(\mathbf{v})_i - \mathbf{S}^M(\mathbf{v})_i}{\mathbf{S}^M(\mathbf{v})_i}$, so that when $\left| \frac{\mathbf{S}^H(\mathbf{v})_i - \mathbf{S}^M(\mathbf{v})_i}{\mathbf{S}^M(\mathbf{v})_i} \right| \leq \beta$, we resort to the high order solution. Also in this case, a too small value of β provides results that are in practice coincident with the low order solution. Appropriate choices for β will also be presented in Section 5.

4 Extension to the Euler equations

In this section we present the extension of the strategy (10) to the Euler equations

$$\frac{\partial \mathbf{w}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{w}) = \mathbf{0}, \quad (11)$$

where

$$\mathbf{w} = \begin{bmatrix} \rho \\ \rho \mathbf{u} \\ \rho E \end{bmatrix} \quad \begin{bmatrix} \rho \mathbf{u} \\ \rho \mathbf{u} \otimes \mathbf{u} + p \mathbf{I} \\ (\rho E + p) \mathbf{u} \end{bmatrix}.$$

Here, ρ is the fluid density, \mathbf{u} is the fluid velocity, p is the pressure and E is the total energy and \mathbf{I} is the d -dimensional identity matrix. The above equations must be complemented by an equation of state (EOS). In this work we consider the classical ideal gas EOS

$$p = (\gamma - 1) \left(\rho E - \frac{1}{2} \rho \mathbf{u} \cdot \mathbf{u} \right), \quad (12)$$

where $\gamma = 1.4$ is the specific heats ratio. Hence, the application of the filtering approach to the density reads as follows:

$$\begin{aligned} \rho_i^F &= \mathbf{S}^M(\boldsymbol{\rho})_i \\ &+ \beta_\rho \mathbf{S}^M(\boldsymbol{\rho})_i F \left(\frac{\mathbf{S}^H(\boldsymbol{\rho})_i - \mathbf{S}^M(\boldsymbol{\rho})_i}{\beta_\rho \mathbf{S}^M(\boldsymbol{\rho})_i} \right), \quad i = 1, \dots, \dim(Q_h), \end{aligned} \quad (13)$$

where $\boldsymbol{\rho}$ is the vector of the degrees of freedom for the density and β_ρ is the tolerance parameter for the density. In analogy to [21], we choose to perform the filtering procedure for all the conserved variables, namely also for $\rho \mathbf{u}$ and ρE , and their formulation is analogous to (13):

$$\begin{aligned} \rho \mathbf{u}_i^F &= \mathbf{S}^M(\rho \mathbf{u})_i \\ &+ \beta_{\rho \mathbf{u}} \mathbf{S}^M(\rho \mathbf{u})_i F \left(\frac{\mathbf{S}^H(\rho \mathbf{u})_i - \mathbf{S}^M(\rho \mathbf{u})_i}{\beta_{\rho \mathbf{u}} \mathbf{S}^M(\rho \mathbf{u})_i} \right), \quad i = 1, \dots, [\dim(Q_h)]^d \end{aligned} \quad (14)$$

$$\begin{aligned} \rho \mathbf{E}_i^F &= \mathbf{S}^M(\rho \mathbf{E})_i \\ &+ \beta_{\rho E} \mathbf{S}^M(\rho \mathbf{E})_i F \left(\frac{\mathbf{S}^H(\rho \mathbf{E})_i - \mathbf{S}^M(\rho \mathbf{E})_i}{\beta_{\rho E} \mathbf{S}^M(\rho \mathbf{E})_i} \right), \quad i = 1, \dots, \dim(Q_h). \end{aligned} \quad (15)$$

5 Numerical experiments

The numerical scheme outlined in the previous Sections has been validated in a number of benchmarks. We set

$$\mathcal{H} = \min\{\text{diam}(K) | K \in \mathcal{T}_h\}$$

and we define the Courant number:

$$C = kU \Delta t / \mathcal{H}, \quad (16)$$

where U is the magnitude of the flow velocity. In the case of the Euler equations, the Courant number C is defined as:

$$C = k(U + c) \Delta t / \mathcal{H}, \quad (17)$$

where $c = \sqrt{\gamma \frac{p}{\rho}}$ is the speed of sound. We chose to employ $k = 1$ and $k = 2$ in combination with the second order SSP and the third order SSP schemes previously recalled in Section 2, respectively.

5.1 Solid body rotation

We consider a classical benchmark for convection schemes, the so-called solid body rotation, which has been studied in different configurations (see e.g. [20], [29]). A stationary velocity field is considered, representing a rotating flow with frequency $\omega = 1 \text{ s}^{-1}$ around the point $(0, 0)$ on the domain $\Omega = (-0.5, 0.5)^2$. The initial datum is given by the following discontinuous function:

$$u_0(\mathbf{x}) = \begin{cases} 1 & \text{if } X^2 + Y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $X = \frac{x-x_0}{\sigma}$ and $Y = \frac{y-y_0}{\sigma}$ with $x_0 = y_0 = \frac{1}{6}$ and $\sigma = 0.2$. For this first test, the computational grid is composed by 120 elements along each direction with a time step such that the maximum Courant number is $C \approx 0.1$. All the results are presented at $T_f = 2\pi$ s, when one rotation has been completed, so that the solution coincides with the initial datum. We first consider the strategy (9) depicted in Section 2 with the filter function $F_1(x)$ in (8), taking $\varepsilon = 5h_K$, as suggested in [5], where $h_K = \text{diam}(K)$. Figure 1 compares the results at $t = T_f$ of the filtering approach with the Q_1 non monotized solution and with the Q_0 one. As one can easily notice, with this choice of the parameter, too much stabilization is added and therefore the filtered solution essentially coincides with the low order one.

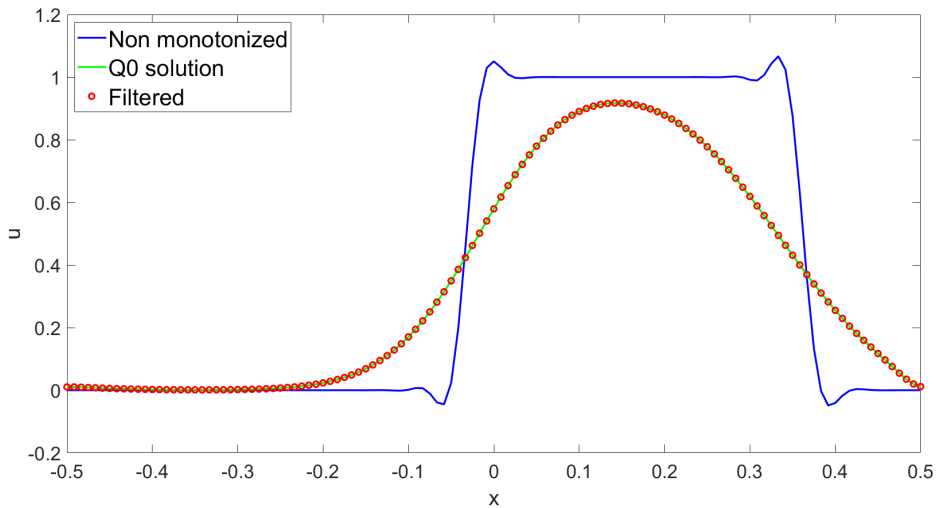


Figure 1: Computational results for the solid body rotation at $t = T_f$ with $k = 1$ using filter (9) with $\varepsilon = 5h_K$. The blue line denotes the non filtered Q_1 solution, the green line denotes the Q_0 solution, while the red dots represent the results of the simulation with the filtering approach.

Increasing the value ε does not affect significantly the results until we take $\varepsilon = 10^4 h_K$: in this case, as evident from Figure 2, the filtering approach works quite well since it is able to provide an essentially monotonic solution, as confirmed by Table 1, without smoothing it too much.

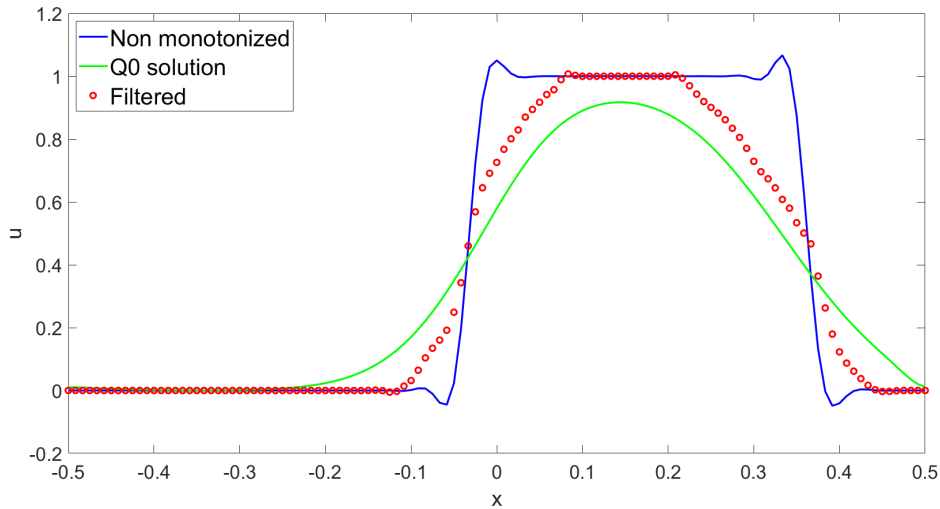


Figure 2: Computational results for the solid body rotation at $t = T_f$ with $k = 1$ using filter (9) with $\varepsilon = 10^4 h_K$. The blue line denotes the non filtered Q_1 solution, the green line denotes the Q_0 solution, while the red dots represent the results of the simulation with the filtering approach.

The situation can be further improved using mesh adaptivity so as to start with a coarse mesh and perform refinement only in the zones where discontinuity is detected. The indicator is based on the gradient of the variable u ; more specifically, we define for each element K

$$\eta_K = \|\nabla u\|_{\infty, K} \quad (18)$$

as local refinement indicator. The initial mesh is composed by 120 elements along each direction and we allowed up to two local refinements. Figure 3 shows that the results at $t = T_f$ with a time step such that the maximum Courant number is $C \approx 0.1$, using the value $\varepsilon = 10^4 h_K$ previously tested in the fixed grid configuration, compared with the full resolution Q_1 non monotized solution and the corresponding Q_0 one. One can easily notice that in this specific configuration the value of ε is still too small and too much dissipation is provided.

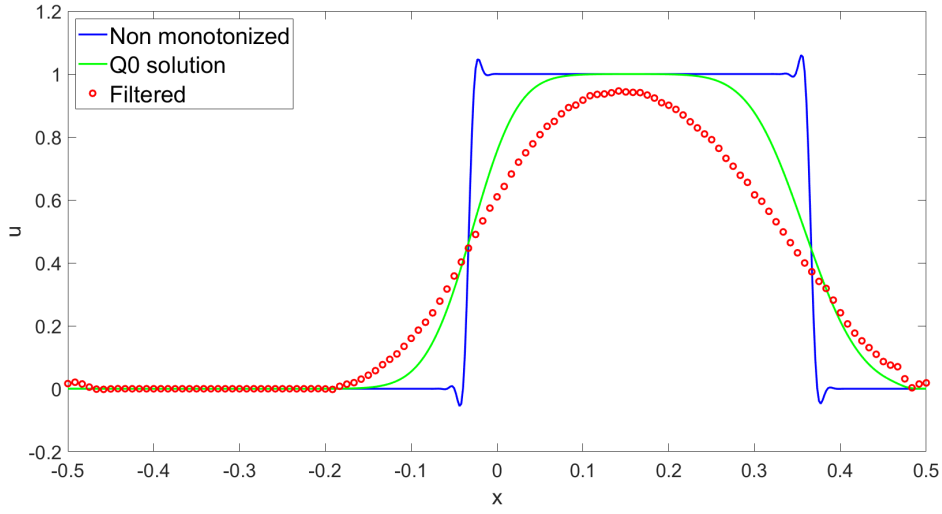


Figure 3: Computational results for adaptive simulation of the solid body rotation at $t = T_f$ with $k = 1$ using filter (9) with $\varepsilon = 10^4 h_K$. The blue line denotes the full resolution non filtered Q_1 solution, the green line denotes the full resolution Q_0 solution, while the red dots represent the results of the simulation with the filtering approach.

The situation improves increasing the value of ε . Figure 4 shows the results using $\varepsilon = 10^5 h_K$, where an essentially monotonic solution is retrieved. The values reported in Table 1 confirm the better quality of the solution.

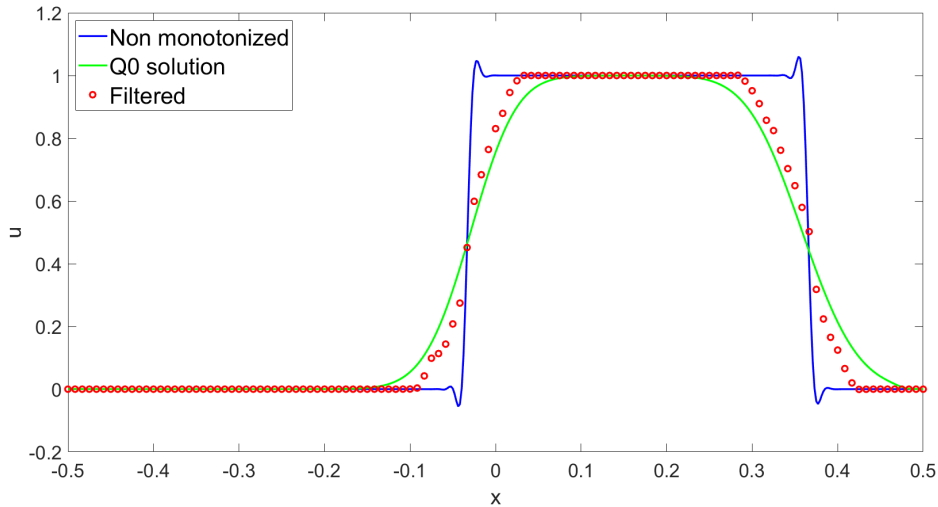


Figure 4: Computational results for adaptive simulation of the solid body rotation at $t = T_f$ using filter (9) with $\varepsilon = 10^5 h_K$. The blue line denotes the full resolution non filtered Q_1 solution, the green line denotes the full resolution Q_0 solution, while the red dots represent the results of the simulation with the filtering approach.

Value of ε	Maximum value	Minimum value
$10^4 h_K$	$1.0 + 7.3 \cdot 10^{-3}$	$0.0 - 5.2 \cdot 10^{-3}$
$10^5 h_K$ (adaptive)	$1.0 + 1.0 \cdot 10^{-5}$	$0.0 - 1.9 \cdot 10^{-3}$

Table 1: Solid body rotation, maximum and minimum values for filtering approach (9) at $t = T_f$ with $k = 1$ both in case of fixed grid and adaptive simulations.

The very large value of ε that is necessary to achieve monotonicity suggests that the previous approach has shortcomings. We consider therefore the second strategy (10) outlined in Section 2. We start again from a fixed grid configuration, using the same mesh and the same time step previously described. After some sensitivity study, $\beta = 0.4$ seems to yield an acceptable behaviour for the solution, as evident from Figure 5. The discontinuity is less smeared out with respect to the Q_0 solution, while avoiding the spurious oscillations and retrieving an essentially monotonic solution, as reported in Table 2.

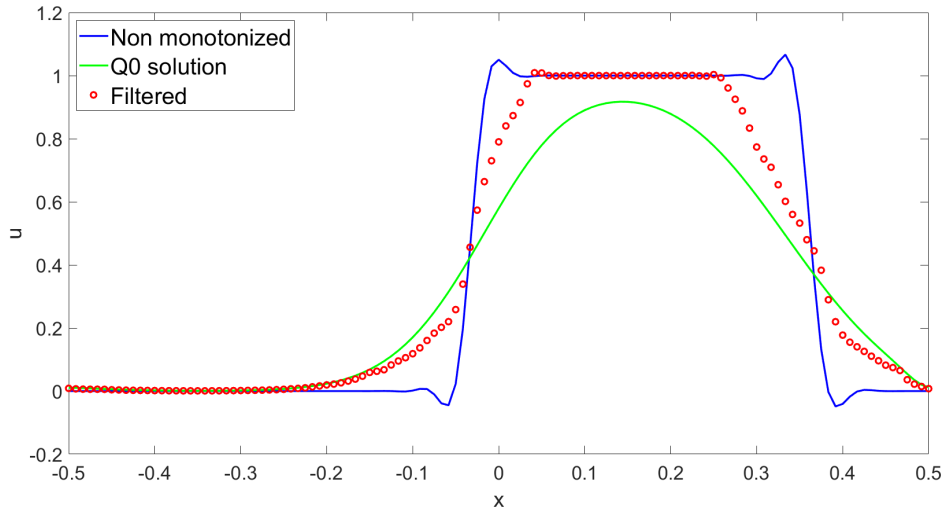


Figure 5: Computational results for the solid body rotation at $t = T_f$ with $k = 1$ using filter (10) with $\beta = 0.4$. The blue line denotes the non filtered Q_1 solution, the green line denotes the Q_0 solution, while the red dots represent the results of the simulation with the filtering approach.

Again, the h -adaptive version of the method, using the same configuration and the same refinement criterion previously described, provides better results, as confirmed by Table 2. The grid at $t = T_f$ is reported in Figure 7 and is composed by 28119 elements.

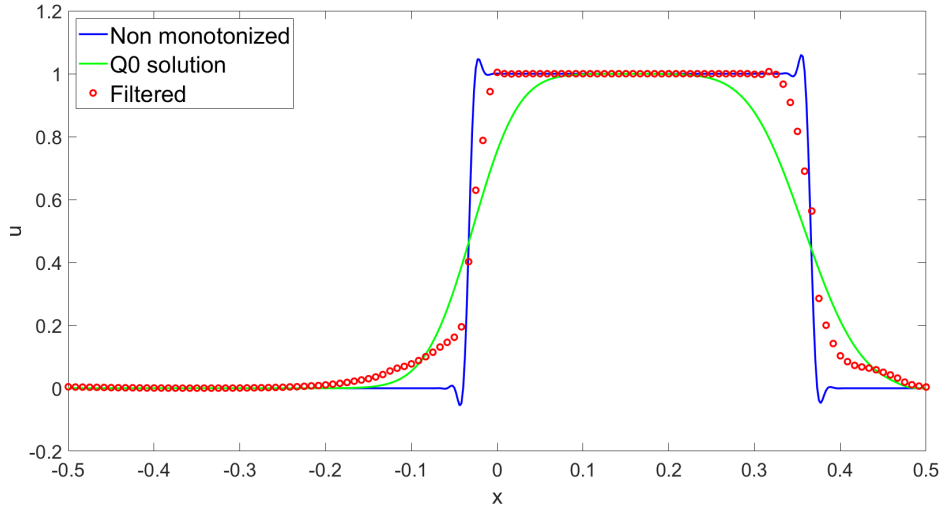


Figure 6: Computational results for adaptive simulation of the solid body rotation at $t = T_f$ with $k = 1$ using filter (10) with $\beta = 0.4$. The blue line denotes the full resolution non filtered Q_1 solution, the green line denotes the full resolution Q_0 solution, while the red dots represent the results of the simulation with the filtering approach.

Value of β	Maximum value	Minimum value
0.4	$1.0 + 9.2 \cdot 10^{-3}$	$0.0 + 7.6 \cdot 10^{-3}$
0.4 (adaptive)	$1.0 + 6.7 \cdot 10^{-3}$	$0.0 + 4.0 \cdot 10^{-4}$

Table 2: Solid body rotation, maximum and minimum values for filtering approach (10) at $t = T_f$ both in case of fixed grid and adaptive simulations.

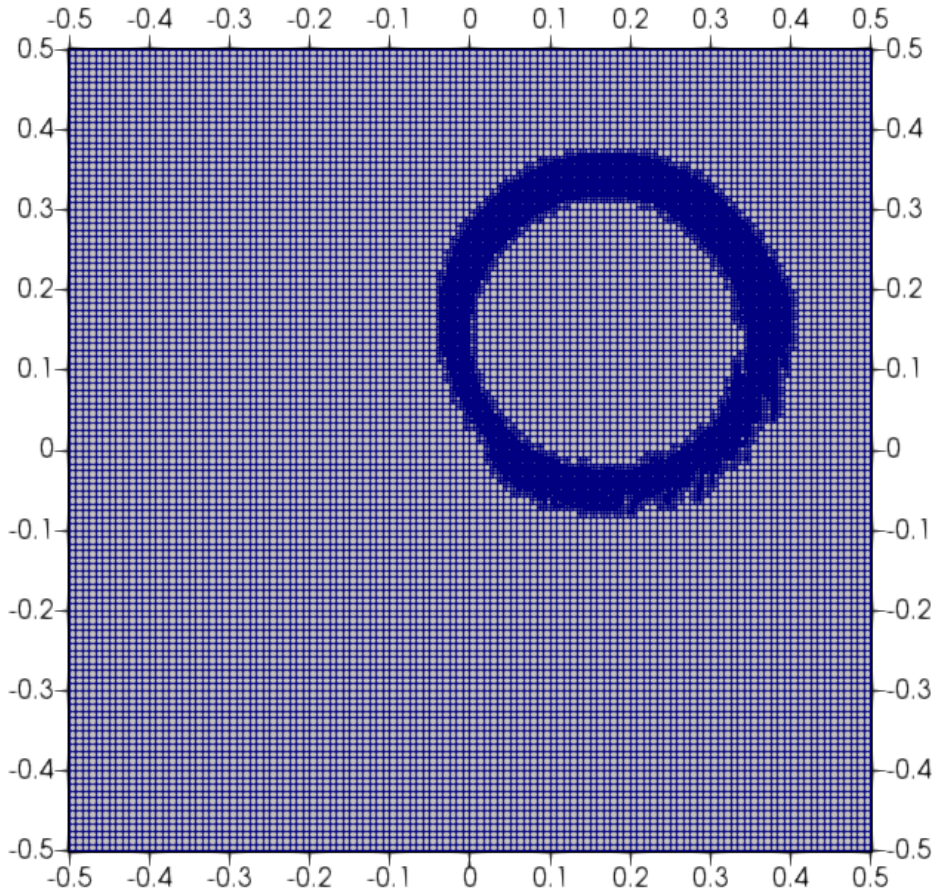


Figure 7: Computational grid for adaptive simulation of the solid body rotation at $t = T_f$ with $k = 1$ using filter (10) with $\beta = 0.4$.

The same test has been repeated using $k = 2$, i.e. Q_2 finite elements, and the third order SSP time discretization strategy briefly recalled in Section 2. We present here only a comparison between the two strategies in case of an adaptive simulation using $\varepsilon = 10^5 h_K$ and $\beta = 0.4$, respectively. Again, we started with a mesh composed by 120 elements along each direction, we allowed up to two local refinements and the employed time step is such that the maximum Courant number is $C \approx 0.1$. Figure 8 shows the results with the two different approaches at $t = T_f$, compared with a full resolution Q_2 solution and the corresponding Q_0 one. One can easily notice that both strategies provide an essentially monotonic result, as confirmed by Table 3; moreover, the approach (10) is characterized by a sharper transition zone and therefore it will be the one used throughout the rest of the numerical experiments.

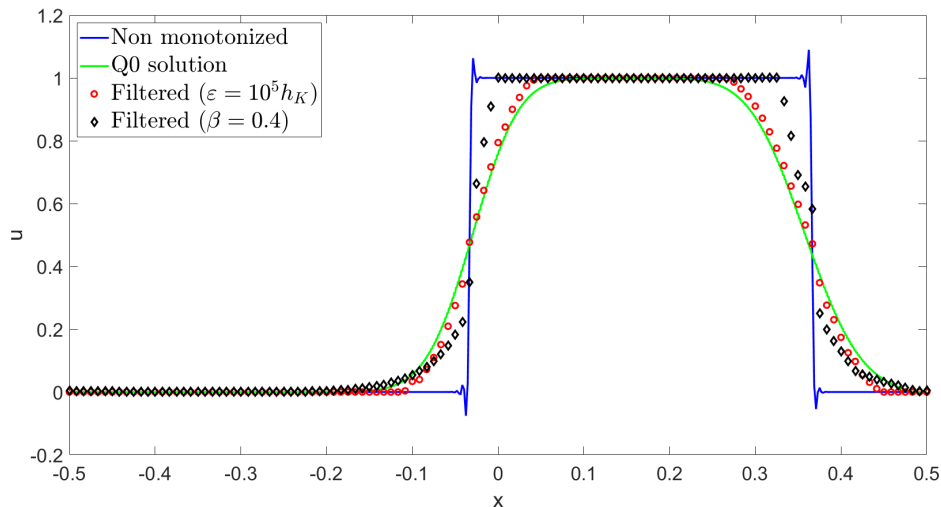


Figure 8: Computational results for adaptive simulations of the solid body rotation at $t = T_f$ with $k = 2$. The blue line denotes the full resolution non filtered Q_2 solution, the green line denotes the full resolution Q_0 solution, while the red dots and black diamonds represent the results of the simulation with the filtering approaches (9) and (10) using $\varepsilon = 10^5 h_K$ and $\beta = 0.4$, respectively.

Value of the parameter β	Maximum value	Minimum value
$\beta = 0.4$ (adaptive)	$1.0 + 1.5 \cdot 10^{-3}$	$0.0 + 3.8 \cdot 10^{-3}$
$\varepsilon = 10^5 h_K$ (adaptive)	1.0	$0.0 - 8.4 \cdot 10^{-5}$

Table 3: Solid body rotation, maximum and minimum values for filtering approach (9) at $t = T_f$ with $k = 2$ in case of adaptive simulations.

5.2 Sod shock tube problem

We consider now the classical Sod shock tube problem proposed by [27] in order to assess the capability of the filtering approach to reproduce correctly 1D waves such as shocks, contact discontinuities or rarefaction waves. It consists of a right-moving shock wave, an intermediate contact discontinuity and a left-moving rarefaction fan. The computational domain is $\Omega = (-0.5, 0.5)$, the final time is $T_f = 0.2$ s and the initial condition is given as follows:

$$(\rho_0, u_0, p_0) = \begin{cases} (1, 0, 1) & \text{if } x < 0 \\ (0.125, 0, 1) & \text{if } x > 0, \end{cases} \quad (19)$$

in terms of density, velocity and pressure, respectively. Dirichlet boundary conditions are imposed. We use as numerical flux the Rusanov [23] flux. We start with a mesh composed by 100 elements and a time-step equal to $5 \cdot 10^{-4}$ s and $k = 1$, yielding a maximum Courant number $C \approx 0.12$. Figure 9 shows the results at $t = T_f$ for the density of a simulation using $\beta_\rho = \beta_{\rho\mathbf{u}} = \beta_{\rho E} = 0.4$. One can easily notice the presence of significant under- and over-shoots. This suggests that we need to decrease the value of the parameter β_ρ in order to achieve a monotonic solution. The same considerations hold also for the velocity and the pressure.

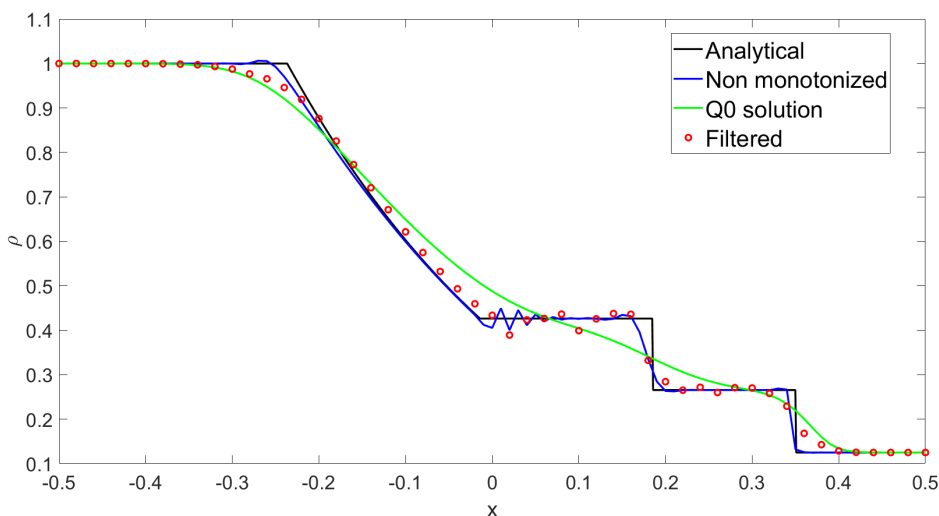


Figure 9: Computational results for Sod shock tube problem at $t = T_f$ with $k = 1$. The black line reports the analytical solution, the blue line denotes the non filtered Q_1 solution, the green line denotes the Q_0 solution, while the red dots represent the results of the simulation with the filtering approach using $\beta_\rho = \beta_{\rho\mathbf{u}} = \beta_{\rho E} = 0.4$.

After some sensitivity study, the combination $\beta_\rho = 0.2, \beta_{\rho\mathbf{u}} = 0.15, \beta_{\rho E} = 0.2$ could be shown to provide a better quality solution with significantly reduced under- and over-shoots, as reported in Figure 10. The situation can be further improved employing the Froese and Oberman's filter function $F_2(x)$ of (8) used in [13], that is continuous and provides therefore a smoother transition between the high order and the low order solutions. This allows also to increase the values of the parameters $\beta_\rho, \beta_{\rho\mathbf{u}}$ and $\beta_{\rho E}$. Figure 11 shows the results at $t = T_f$ using $\beta_\rho = \beta_{\rho\mathbf{u}} = \beta_{\rho E} = 0.3$ and one can easily notice that the shock wave and the contact discontinuity are resolved in a sharper manner with only slight undershoots for density and pressure

and overshoots for the velocity in the tail of the rarefaction fan. Table 4 reports the maximum and the minimum values for density, velocity and pressure, as well as the L^∞ norm errors, which confirm the good results of the proposed method, also in comparison with the results obtained in [21] with the classical MOOD scheme.

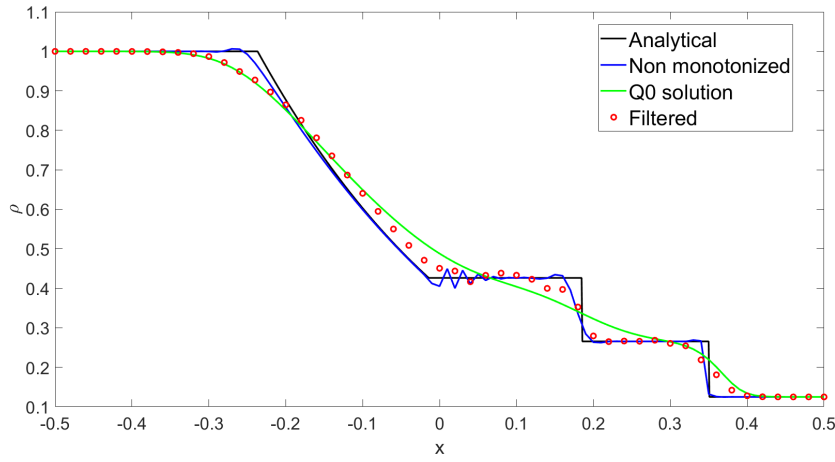
Variable	Maximum value	Minimum value	L^∞ error	L^∞ error [21]
ρ	1.0	0.125	$6.9 \cdot 10^{-2}$	$1.1 \cdot 10^{-1}$
u	$0.9275 + 5.0 \cdot 10^{-2}$	0.0	$4.8 \cdot 10^{-1}$	
p	1.0	0.1	$8.2 \cdot 10^{-2}$	

Table 4: Computational results for Sod shock tube problem at $t = T_f$ with $k = 1$ employing Froese and Oberman’s filter function and using $\beta_\rho = 0.3$, $\beta_{\rho\mathbf{u}} = 0.3$ and $\beta_{\rho E} = 0.3$.

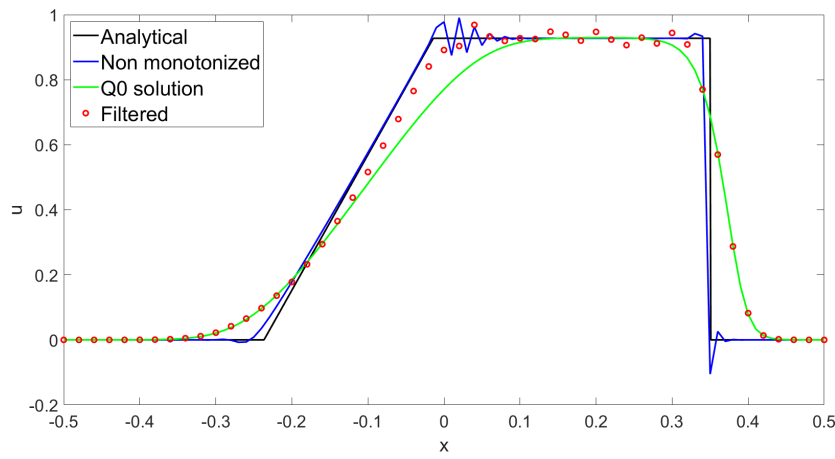
Figure 12 reports the results at $t = T_f$ using 250 elements, a time step equal to $2 \cdot 10^{-4}$ s and the following parameters: $\beta_\rho = 0.6$, $\beta_{\rho\mathbf{u}} = 0.6$, $\beta_{\rho E} = 0.6$. It can be easily noticed that, as expected by increasing the resolution, the discontinuities are better retrieved. The values reported in Table 5 confirm the improved results.

Variable	Maximum value	Minimum value	L^∞ error
ρ	1.0	0.125	$3.9 \cdot 10^{-2}$
u	$0.9275 + 4.2 \cdot 10^{-2}$	0.0	$2.9 \cdot 10^{-1}$
p	1.0	0.1	$5.0 \cdot 10^{-2}$

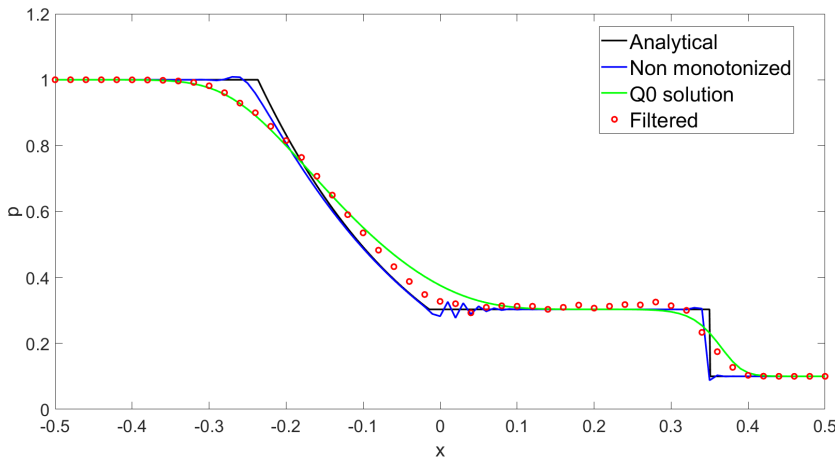
Table 5: Computational results for Sod shock tube problem at $t = T_f$ with $k = 1$ and 250 elements employing Froese and Oberman’s filter function and using $\beta_\rho = 0.6$, $\beta_{\rho\mathbf{u}} = 0.6$ and $\beta_{\rho E} = 0.6$.



a)

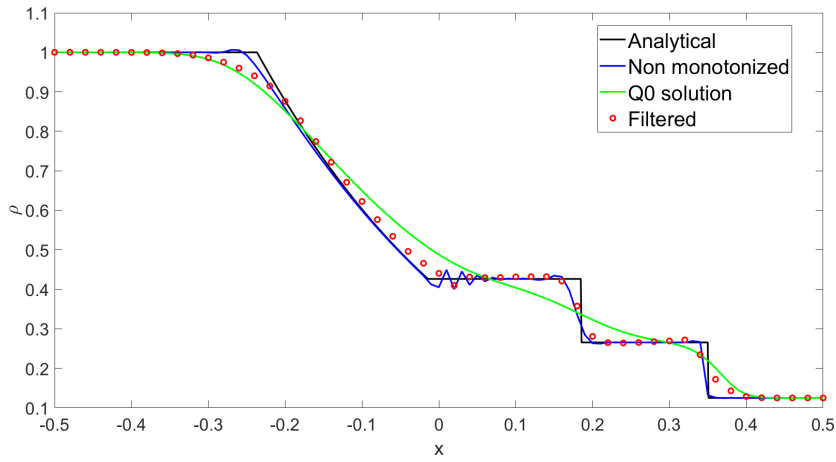


b)

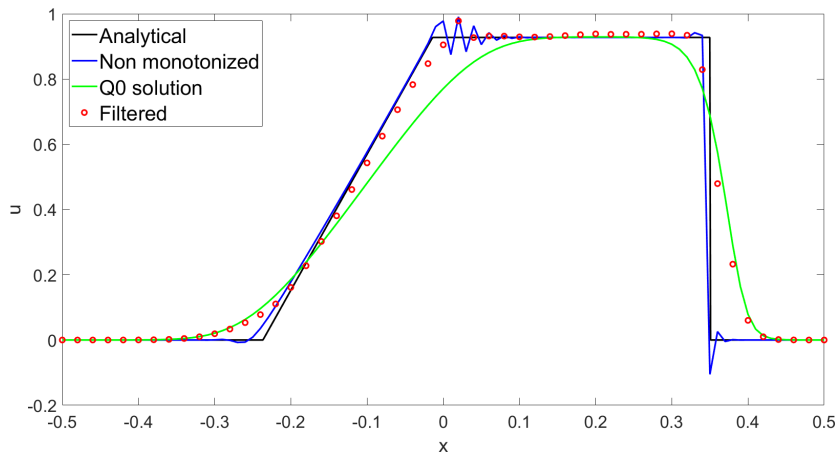


c)

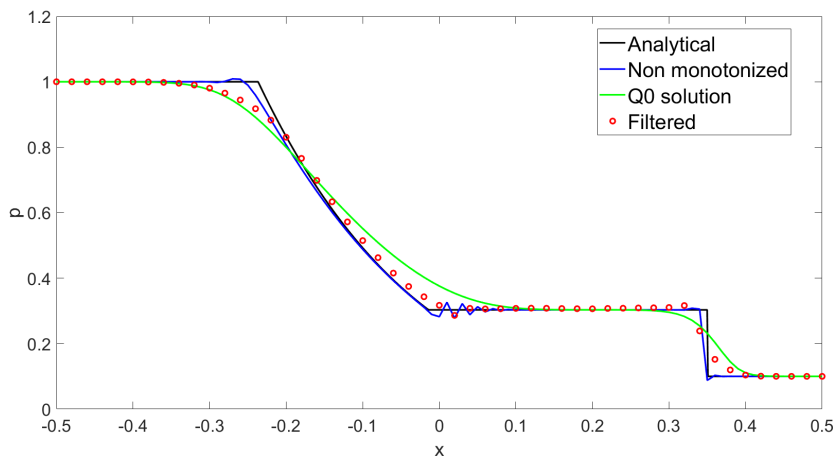
Figure 10: Computational results for Sod shock tube problem at $t = T_f$ with $k = 1$, a) density, b) velocity, c) pressure. The black line reports the analytical solution, the blue line denotes the non filtered Q_1 solution, the green line denotes the Q_0 solution, while the red dots represent the results of the simulation with the filtering approach using $\beta_\rho = 0.2$, $\beta_{\rho u} = 0.15$ and $\beta_{\rho E} = 0.2$.



a)

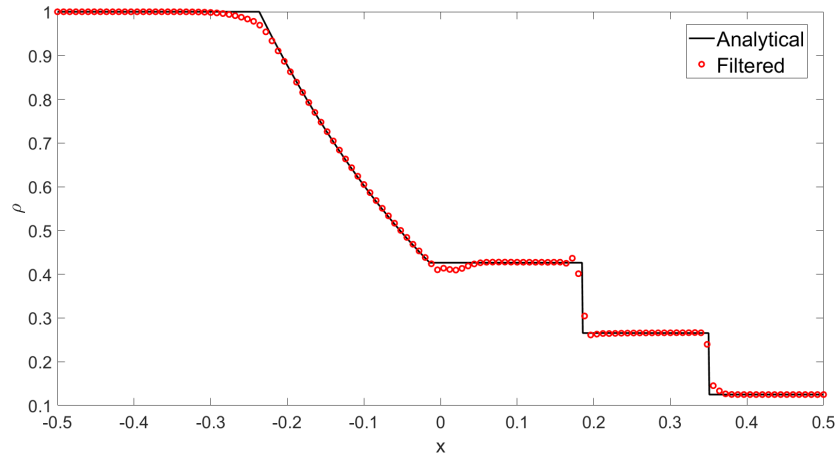


b)

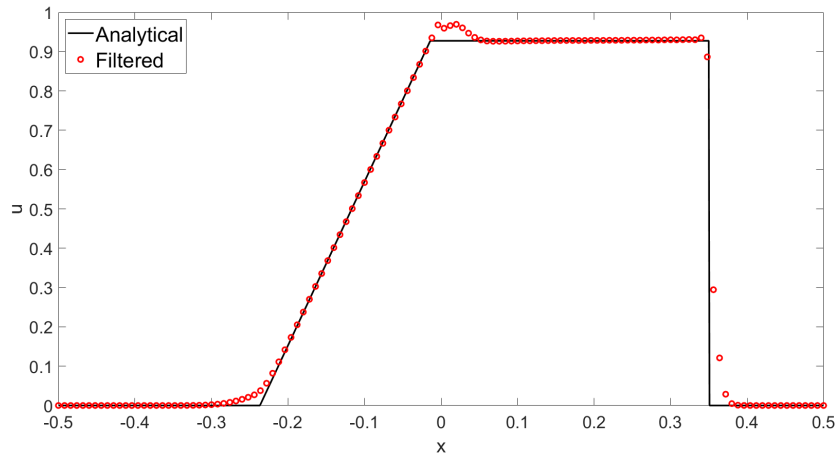


c)

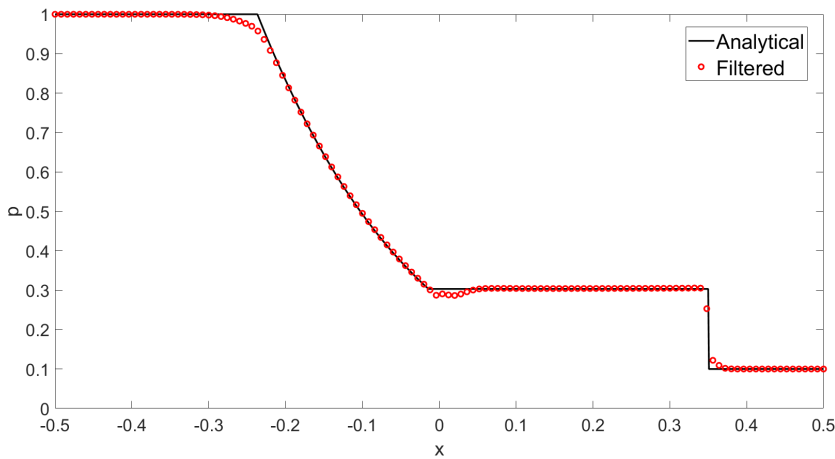
Figure 11: Computational results for Sod shock tube problem at $t = T_f$ with $k = 1$, a) density, b) velocity, c) pressure. The black line reports the analytical solution, the blue line denotes the non filtered Q_1 solution, the green line denotes the Q_0 solution, while the red dots represent the results employing Froese and Oberman's filter function and using $\beta_\rho = 0.3$, $\beta_{\rho u} = 0.3$ and $\beta_{\rho E} = 0.3$.



a)



b)



c)

Figure 12: Computational results for Sod shock tube problem at $t = T_f$ with $k = 1$, a) density, b) velocity, c) pressure. The black line reports the analytical solution, while the red dots represent the results of the simulation with 250 elements employing Froese and Oberman's filter function and using $\beta_\rho = 0.6$, $\beta_{\rho u} = 0.6$ and $\beta_{\rho E} = 0.6$.

The same test has been repeated using $k = 2$ and the third order SSP time discretization scheme. Figure 13 reports the results at $t = T_f$ using 250 elements, a time step equal to $1 \cdot 10^{-4}$ s and $\beta_\rho = \beta_{\rho\mathbf{u}} = \beta_{\rho E} = 1.4$. The under- and over-shoots are significantly reduced and a good agreement with the analytical solution is established. The larger values of β parameters can be explained by considering that the increase of the polynomial degree leads generally to a more accurate solution with relatively large under- and over-shoots localized in a narrow region, where the low order solution has to be considered.

Variable	Maximum value	Minimum value	L^∞ error
ρ	$1.0 + 2.0 \cdot 10^{-3}$	$0.125 - 2.0 \cdot 10^{-4}$	$1.6 \cdot 10^{-2}$
u	$0.9275 + 8.0 \cdot 10^{-3}$	$0.0 - 2.3 \cdot 10^{-3}$	$2.2 \cdot 10^{-2}$
p	$1.0 + 2.8 \cdot 10^{-3}$	$0.1 - 2.0 \cdot 10^{-4}$	$1.3 \cdot 10^{-2}$

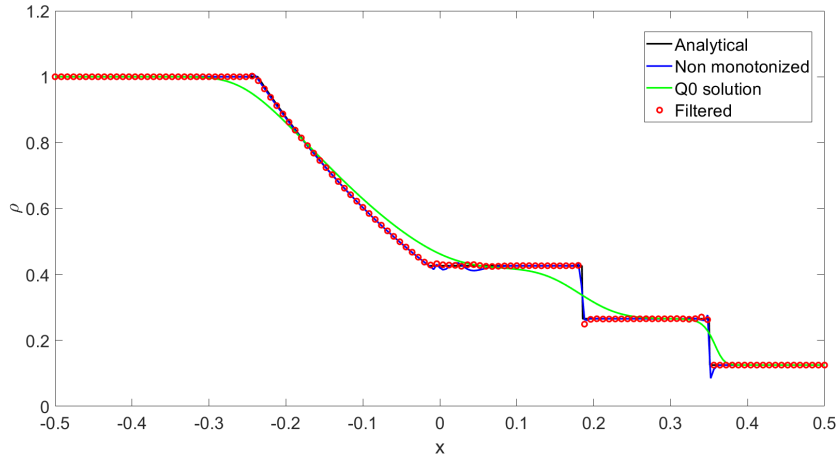
Table 6: Computational results for Sod shock tube problem at $t = T_f$ with $k = 2$ employing Froese and Oberman's filter function and using $\beta_\rho = 1.4$, $\beta_{\rho\mathbf{u}} = 1.4$ and $\beta_{\rho E} = 1.4$.

5.3 Circular explosion problem

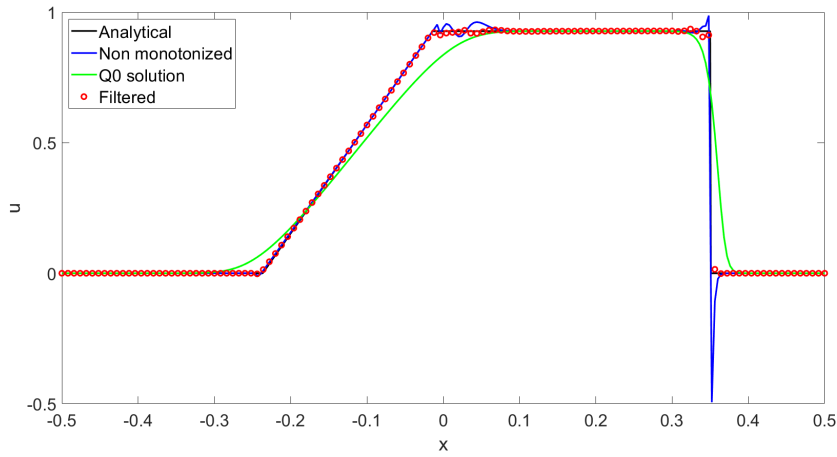
In this section, we consider the two-dimensional explosion problem discussed in [12, 30]. This test is quite relevant since it involves the propagation of waves that are not aligned with the mesh and therefore it can be used to check the ability of the proposed method to preserve physical symmetries of the problem as well as to validate it in multiple space dimensions. The computational domain is $\Omega = (-1, 1)^2$, the final time is $T_f = 0.2$ s and the initial condition is the following:

$$(\rho_0, u_0, v_0, p_0) = \begin{cases} (1, 0, 0, 1) & \text{if } r \leq R \\ (0.125, 0, 0, 1) & \text{if } r > R, \end{cases} \quad (20)$$

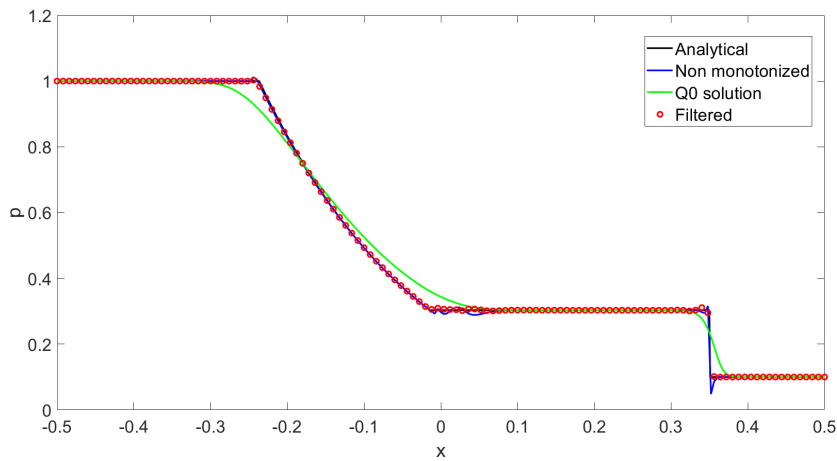
with $R = 0.5$ denoting the radius of initial discontinuity and $r = \sqrt{x^2 + y^2}$ representing the radial distance. As explained in [28], in 2D we have cylindrical symmetry and a reference solution can be computed solving a one dimensional problem in the radial direction with suitable geometric source terms. Figure 14 shows the results obtained using $N_{el} = 200$ elements along each direction, $k = 1$ and $\beta_\rho = \beta_{\rho\mathbf{u}} = \beta_{\rho E} = 1$. One can easily notice that the discontinuities are well reproduced, even using only first order degree polynomial for the high order method, and their position is well captured with only slight undershoots and overshoots in correspondence of the rarefaction wave.



a)



b)



c)

Figure 13: Computational results for Sod shock tube problem at $t = T_f$ with $k = 2$, a) density, b) velocity, c) pressure. The black line reports the analytical solution, the blue line denotes the non filtered Q_2 solution, the green line denotes the Q_0 solution, while the red dots represent the results of the simulation with Froese and Oberman's filter function using $\beta_\rho = 1.4$, $\beta_{\rho\mathbf{u}} = 1.4$ and $\beta_{\rho E} = 1.4$.

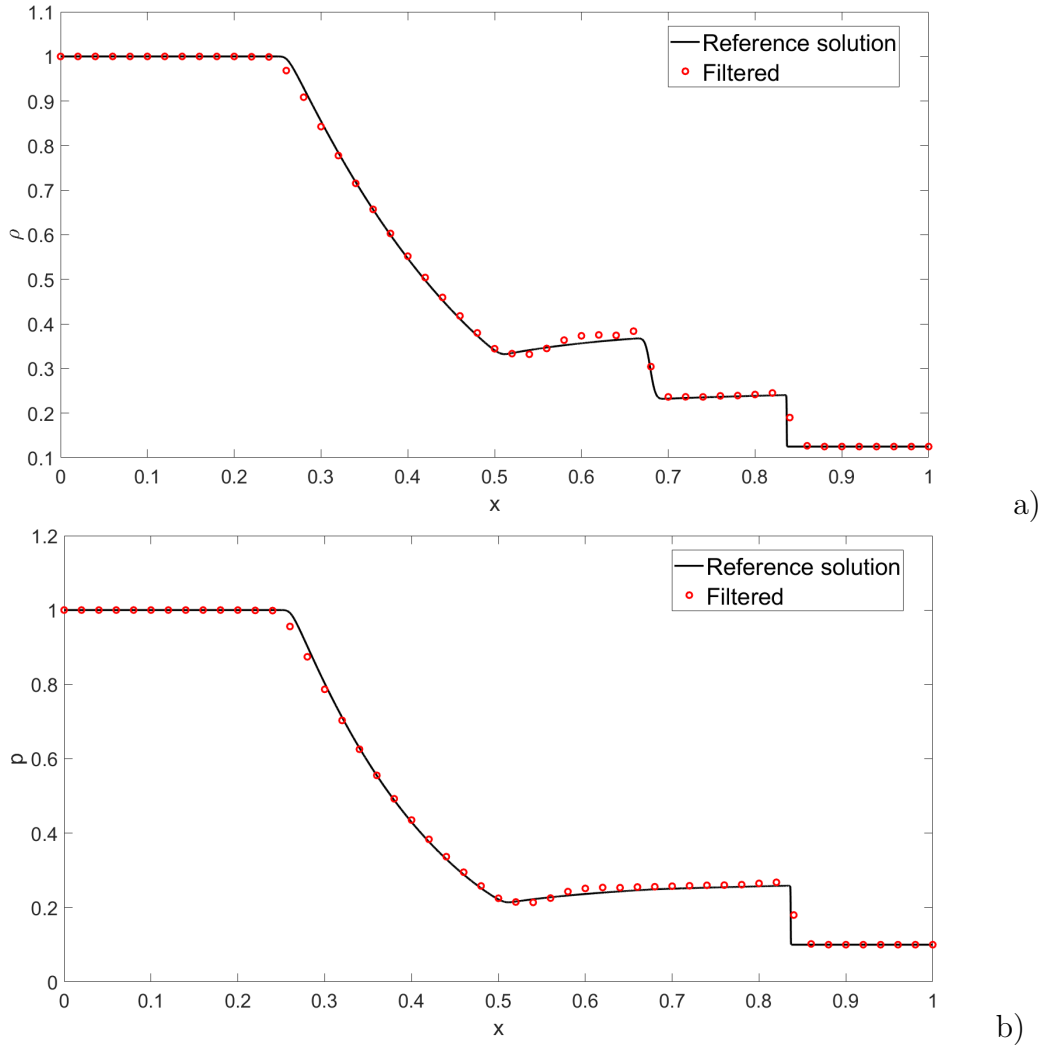


Figure 14: Computational results for 2D explosion problem at $t = T_f$ with $k = 1$, a) density, b) pressure. The black line reports the reference solution computed solving the 1D problem in the radial direction, while the red dots represent the results of the simulation with Froese and Oberman's filter function using $\beta_\rho = 1$, $\beta_{\rho\mathbf{u}} = 1$ and $\beta_{\rho E} = 1$.

The same test has been repeated increasing both the spatial resolution with $N_{el} = 400$ and the high order polynomial degree with $k = 2$. Figure 15 reports the results obtained using $\beta_\rho = \beta_{\rho\mathbf{u}} = \beta_{\rho E} = 1.7$ and an excellent agreement with the reference solution is achieved. Analogous results have been obtained in [30], where however polynomials of degree 9 were employed.

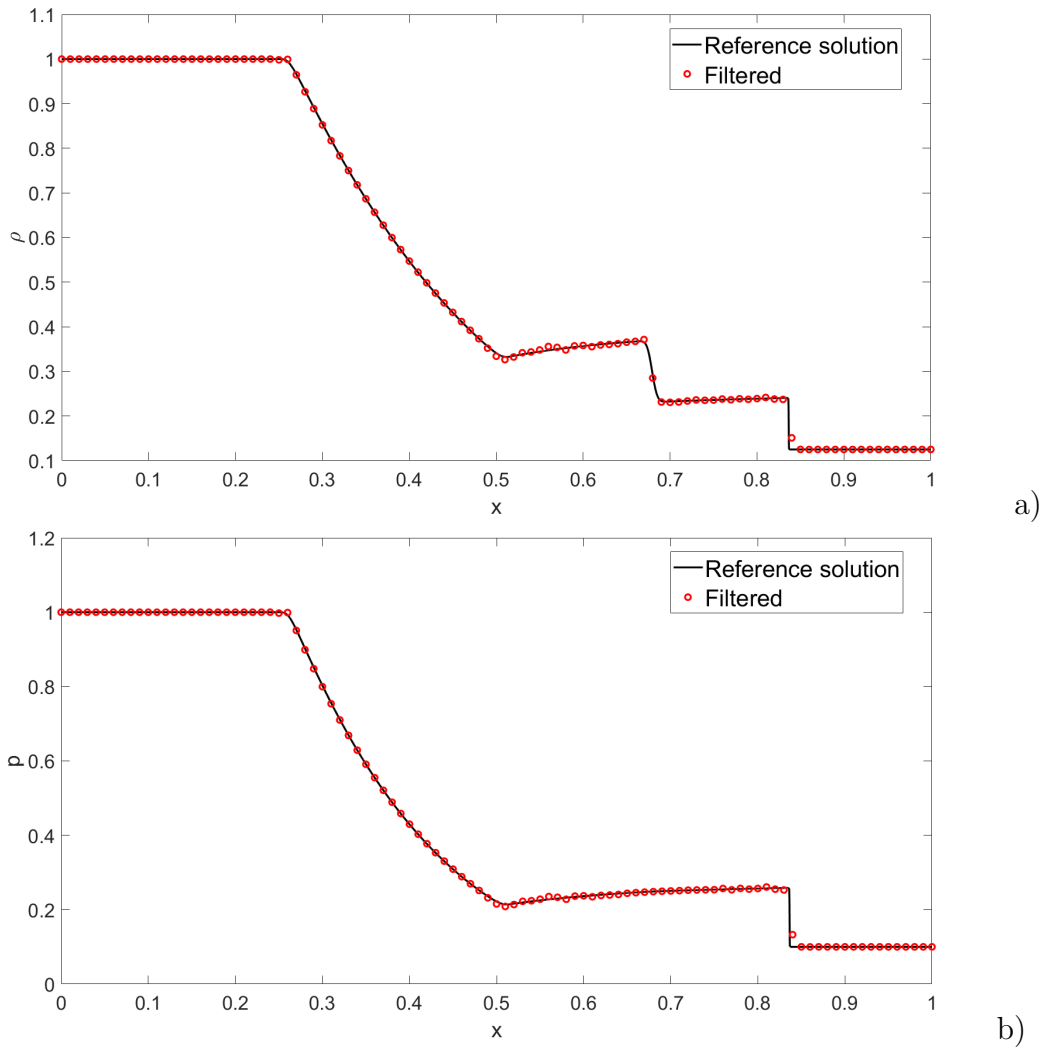


Figure 15: Computational results for 2D explosion problem at $t = T_f$ with $k = 2$, a) density, b) pressure. The black line reports the reference solution computed solving the 1D problem in the radial direction, while the red dots represent the results of the simulation with Froese and Oberman's filter function using $\beta_\rho = 1.7$, $\beta_{\rho u} = 1.7$ and $\beta_{\rho E} = 1.7$.

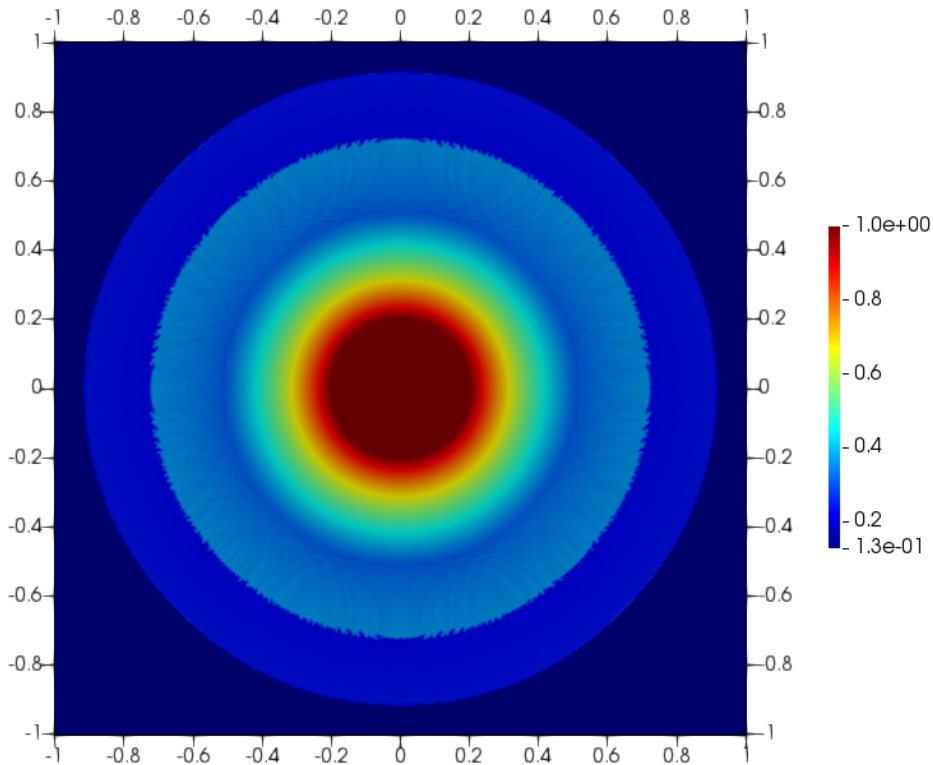


Figure 16: Computational results for 2D explosion problem at $t = T_f$ with $k = 2$, contour plot of the density.

Finally, we have employed the h -adaptive version of the method, starting from a coarse mesh with $N_{el} = 200$ elements along each direction and allowing up to three local refinements which would correspond to a uniform grid with $N_{el} = 1600$. The employed local indicator is based on the gradient of the density; more specifically we define for each element K

$$\eta_K = \|\|\nabla\rho\|\|_{\infty,K}. \quad (21)$$

Figure 18 shows the final grid obtained at $t = T_f$ composed by 63136 elements and one can easily notice that more resolution is added in correspondence of the discontinuities.

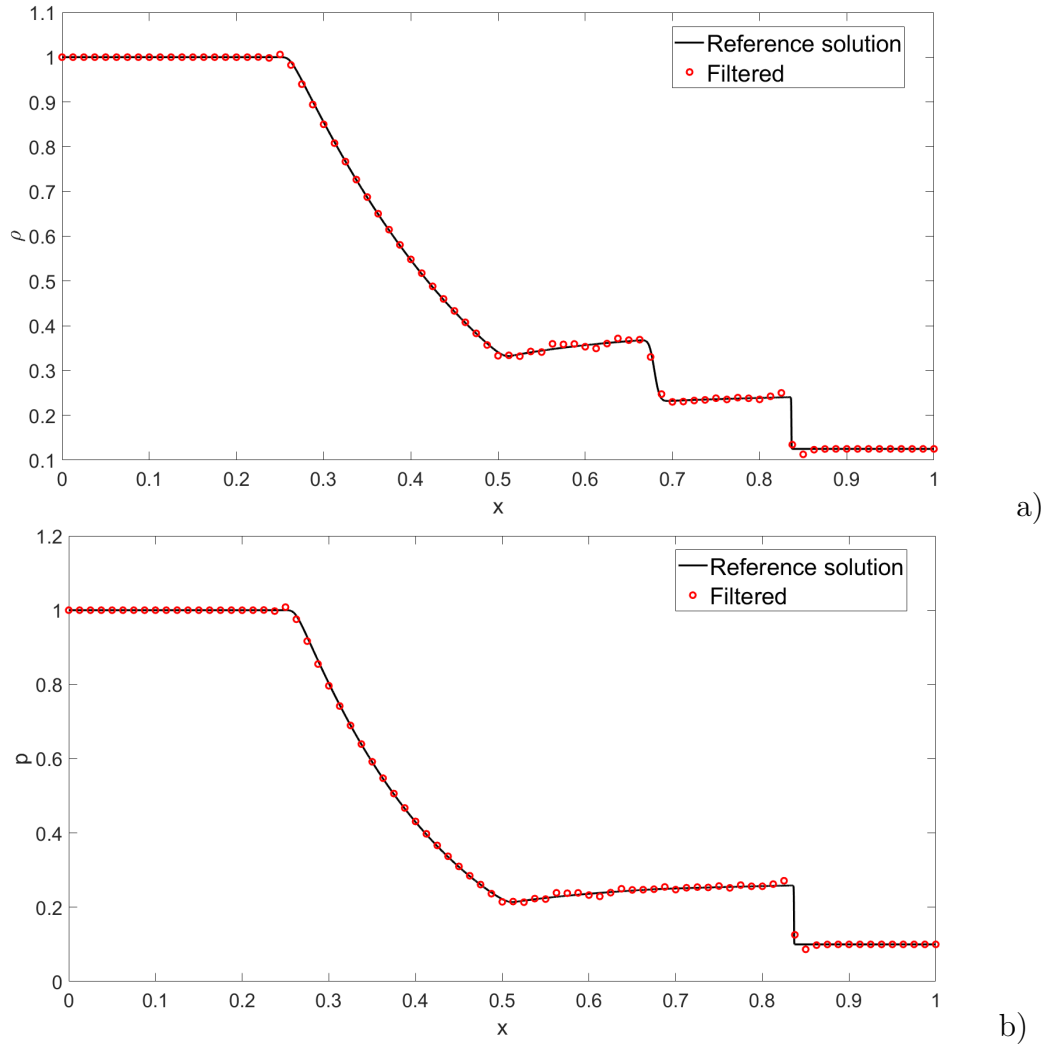


Figure 17: Computational results for 2D explosion problem at $t = T_f$ with adaptive grid and $k = 2$, a) density, b) pressure. The black line reports the reference solution computed solving the 1D problem in the radial direction, while the red dots represent the results of the simulation with Froese and Oberman's filter function using $\beta_\rho = 1.7$, $\beta_{\rho\mathbf{u}} = 1.7$ and $\beta_{\rho E} = 1.7$.

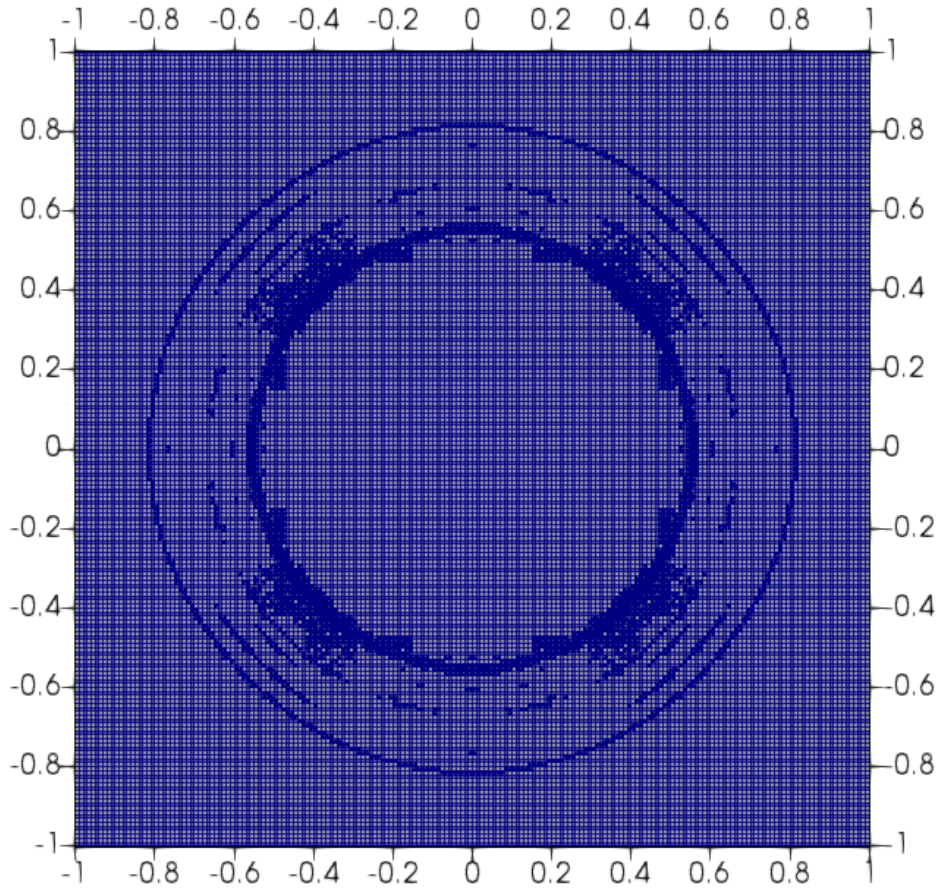


Figure 18: 2D explosion problem, adaptive grid at $t = T_f$ with $k = 2$ and Froese and Oberman's filter function using $\beta_\rho = 1.7$, $\beta_{\rho\mathbf{u}} = 1.7$ and $\beta_{\rho E} = 1.7$.

6 Conclusions and future developments

In this work, a filtering technique for obtaining a monotonic Discontinuous Galerkin discretization of hyperbolic equations has been presented. The scheme is inspired by the approach originally proposed in [5] and it is based on a filter function that keeps the high order solution if it is regular and switches to a monotone low order approximation otherwise, according to the value of one or more parameters. Its potential has been demonstrated in a number of classical benchmarks for linear advection and Euler equations.

In future work, an obvious and necessary development concerns the tuning of the parameter(s) β . The goal is to automatically choose suitable values depending on the employed time and space steps as well as the polynomial degree used by the higher order discretiza-

tion. Moreover, we plan to investigate the behaviour of the proposed method in case stiff source terms and/or non-conservative terms are present, as for example in the Baer-Nunziato model of compressible multiphase flows [1].

Acknowledgements

The author would like to thank Luca Bonaventura for several useful discussions.

References

- [1] M. R. Baer and J. W. Nunziato. A two-phase mixture theory for the deflagration-to-detonation transition (ddt) in reactive granular materials. *International Journal of Multiphase Flow*, 12:861–889, 1986.
- [2] W. Bangerth, R. Hartmann, and G. Kanschat. deal II: a general-purpose object-oriented finite element library. *ACM Transactions on Mathematical Software (TOMS)*, 33:24–51, 2007.
- [3] F. Bassi and S. Rebay. High-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *Journal of Computational Physics*, 131:267–279, 1997.
- [4] F. Bassi and S. Rebay. High-order accurate discontinuous finite element solution of the 2d Euler equations. *Journal of Computational Physics*, 138:251–285, 1997.
- [5] O. Bokanowski, M. Falcone, and S. Sahu. An efficient filtered scheme for some first order time-dependent Hamilton–Jacobi equations. *SIAM Journal on Scientific Computing*, 38:A171–A195, 01 2016.
- [6] B. Cockburn, S. Hou, and C.W. Shu. The Runge-Kutta Local Projection Galerkin Finite Element Method for conservation laws IV: the multidimensional case. *Mathematics of Computation*, 54(190):545–581, 1990.
- [7] B. Cockburn, S.Y. Lin, and C.W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems. *Journal of Computational Physics*, 84(1):90–113, 1989.

- [8] B. Cockburn and C.W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Mathematics of Computation*, 52(186):411–435, 1989.
- [9] B. Cockburn and C.W. Shu. The Runge-Kutta local projection P1 Discontinuous Galerkin method for scalar conservation laws. *Mathematical Modelling and Numerical Analysis*, 25:337–361, 1991.
- [10] B. Cockburn and C.W. Shu. The Runge-Kutta Discontinuous Galerkin method for conservation laws, V. *Journal of Computational Physics*, 141:198–224, 1998.
- [11] M. Dumbser and R. Loubère. A simple robust and accurate a posteriori sub-cell finite volume limiter for the discontinuous Galerkin method on unstructured meshes. *Journal of Computational Physics*, 319:163–199, 2016.
- [12] M. Dumbser, O. Zanotti, R. Loubère, and S. Diot. A posteriori subcell limiting of the discontinuous galerkin finite element method for hyperbolic conservation laws. *Journal of Computational Physics*, 278:47–75, 2014.
- [13] B. D. Froese and A. M. Oberman. Convergent filtered schemes for the Monge-Ampère partial differential equation. *SIAM J. Numer. Anal.*, 51:423–444, 2013.
- [14] F.X. Giraldo. *An Introduction to Element-Based Galerkin Methods on Tensor-Product Bases*. Springer Nature, 2020.
- [15] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of Computation*, 67(221):73–85, 1998.
- [16] S. Gottlieb, C. W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43, 05 2001.
- [17] G.E. Karniadakis and S. Sherwin. *Spectral hp–Element Methods for Computational Fluid Dynamics*. Oxford University Press, 2005.
- [18] D. Kuzmin, R. Löhner, and S. Turek. *Flux-corrected transport: principles, algorithms, and applications*. Springer Verlag, 2012.
- [19] D. Kuzmin and S. Turek. Flux correction tools for finite elements. *Journal of Computational Physics*, 175:525–558, 2002.
- [20] R. J. LeVeque. High-resolution conservative algorithms for advection in incompressible flow. *SIAM Journal on Numerical Analysis*, 33:627–665, 1996.

- [21] R. Loubère, M. Dumbser, and S. Diot. A new family of high order unstructured MOOD and ADER finite volume schemes for multidimensional systems of hyperbolic conservation laws. *Communications in Computational Physics*, 16:718–763, 2014.
- [22] M. Restelli, L. Bonaventura, and R. Sacco. A semi-Lagrangian Discontinuous Galerkin method for scalar advection by incompressible flows. *Journal of Computational Physics*, 216:195–215, 2006.
- [23] V. Rusanov. The calculation of the interaction of non-stationary shock waves and obstacles. *Ussr Computational Mathematics and Mathematical Physics*, 1:304–320, 1962.
- [24] S. Sahu. *High-order and coupled schemes for Hamilton-Jacobi-Bellman equations*. PhD thesis, Università Sapienza, Roma, Italy, 2015.
- [25] C.W. Shu. High-order finite difference and finite volume WENO schemes and discontinuous Galerkin methods for CFD. *International Journal of Computational Fluid Dynamics*, 17:107–118, 2003.
- [26] C.W. Shu. High order WENO and DG methods for time-dependent convection-dominated PDEs: A brief survey of several recent developments. *Journal of Computational Physics*, 316:598–613, 2016.
- [27] G.A. Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics*, 27(1):1–31, 1978.
- [28] E. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer, 01 2009.
- [29] S. T. Zalesak. Fully multidimensional flux-corrected transport algorithms for fluids. *Journal of Computational Physics*, 31:335–362, 1979.
- [30] O. Zanotti, F. Fambri, M. Dumbser, and A. Hidalgo. Space-time adaptive ADER discontinuous Galerkin finite element schemes with a posteriori sub-cell finite volume limiting. *Computers & Fluids*, 118:204–224, 2015.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 24/2022** Cappozzo, A.; McCrory, C.; Robinson, O.; Freni Sterrantino, A.; Sacerdote, C.; Krogh, V.; Pan
A blood DNA methylation biomarker for predicting short-term risk of cardiovascular events
- 25/2022** Cavinato, L.; Gozzi, N.; Sollini, M.; Kirienko, M.; Carlo-Stella, C.; Rusconi, C.; Chiti, A.; Ieva, F.
Perspective transfer model building via imaging-based rules extraction from retrospective cancer subtyping in Hodgkin Lymphoma
- 23/2022** Masci, C.; Ieva, F.; Paganoni, A.M.
A multinomial mixed-effects model with discrete random effects for modelling dependence across response categories
- 22/2022** Regazzoni, F.; Pagani, S.; Quarteroni, A.
Universal Solution Manifold Networks (USM-Nets): non-intrusive mesh-free surrogate models for problems in variable domains
- 21/2022** Cappozzo, A.; Ieva, F.; Fiorito, G.
A general framework for penalized mixed-effects multitask learning with applications on DNA methylation surrogate biomarkers creation
- 20/2022** Clementi, L.; Gregorio, C.; Savarè, L.; Ieva, F.; Santambrogio, M.D.; Sangalli, L.M.
A Functional Data Analysis Approach to Left Ventricular Remodeling Assessment
- 19/2022** Lupo Pasini, M.; Perotto, S.
Hierarchical model reduction driven by machine learning for parametric advection-diffusion-reaction problems in the presence of noisy data
- 18/2022** Bennati, L.; Vergara, C.; Giamb Bruno, V.; Fumagalli, I.; Corno, A.F.; Quarteroni, A.; Puppini, G.; L
An image-based computational fluid dynamics study of mitral regurgitation in presence of prolapse
- 17/2022** Regazzoni, F.
Stabilization of staggered time discretization schemes for 0D-3D fluid-structure interaction problems
- 16/2022** G. Ciaramella, T. Vanzan
Substructured Two-grid and Multi-grid Domain Decomposition Methods