TVD-MOOD schemes based on implicit explicit time integration

Victor Michel-Dansac, Andrea Thomann

TVD-MOOD schemes based on implicit explicit time integration

Victor Michel-Dansac[a], Andrea Thomann[b,*]

[a] *Université de Strasbourg, CNRS, Inria, IRMA, F-67000 Strasbourg, France*
[b] *Institut für Mathematik, Johannes Gutenberg-Universität Mainz, Germany*

Abstract

The context of this work is the development of first order total variation diminishing (TVD) implicit-explicit (IMEX) Runge-Kutta (RK) schemes as a basis of a Multidimensional Optimal Order detection (MOOD) approach to approximate the solution of hyperbolic multi-scale equations. A key feature of our newly proposed TVD schemes is that the resulting CFL condition does not depend on the fast waves of the considered model, as long as they are integrated implicitly. However, a result from Gottlieb et al. [18] gives a first order barrier for unconditionally stable implicit TVD-RK schemes and TVD-IMEX-RK schemes with scale-independent CFL conditions. Therefore, the goal of this work is to consistently improve the resolution of a first-order IMEX-RK scheme, while retaining its L^∞ stability and TVD properties. In this work we present a novel approach based on a convex combination between a first-order TVD IMEX Euler scheme and a potentially oscillatory high-order IMEX-RK scheme. We derive and analyse the TVD property for a scalar multi-scale equation and numerically assess the performance of our TVD schemes compared to standard L-stable and SSP IMEX RK schemes from the literature. Finally, the resulting TVD-MOOD schemes are applied to the isentropic Euler equations.

Keywords: MOOD, L^∞ stability, TVD schemes, IMEX RK schemes, isentropic Euler equations

1. Introduction

Multi-scale equations arise in a wide range of applications, such as shallow flows [3], magnetohydrodynamics [30], multi-material interfaces [1] or atmospheric flows [27]. When developing numerical methods for such applications, it is of prime importance to obtain physically admissible solutions under these multi-scale constraints. In order to numerically treat these different scales, one must assess whether the fast scales are relevant to the physical solution. Indeed, accurately capturing these fast scales requires a very restrictive time step. This issue is discussed e.g. in [19] for the Euler equations. When the impact of the fast scales on the physical solution is less important, numerical methods that do not accurately capture all scales, but only follow the slow dynamics, are necessary. One option, which we will study in this paper, is to use Implicit-Explicit (IMEX) schemes, where the terms associated to the fast wave propagation are treated implicitly. Such schemes are well-studied in the literature, see for instance [2] for efficient IMEX schemes applied to hyperbolic-parabolic problems, [35] for IMEX schemes adapted to stiff relaxation source terms, or [33, 12, 4], and references therein, for IMEX schemes designed for the low Mach regime of the Euler equations. In this work, we are concerned with hyperbolic systems whose stiffness originates from the flux, rather than a source term. Let us emphasise that we will not consider hyperbolic systems with stiff source terms typically arising from relaxation processes. For their treatment, we refer for instance to [35].

To increase the quality of numerical approximations, one may turn to high-order schemes. However, such schemes are known to introduce spurious oscillations in the solution away from smooth regions. This is problematic, especially when considering non-linear hyperbolic equations, as the solution can develop discontinuities even when starting with a smooth initial condition. This was already observed by Harten

*Corresponding author

in [20], who introduced the notion of total variation diminishing (TVD) schemes, and constructed non-oscillatory explicit and implicit second-order TVD schemes. Those schemes are non-linear, even when applied on linear equations, as from Godunov's theorem [13] it follows that linear TVD schemes can only be first-order accurate. Since non-linear implicit schemes are computationally very costly, especially when applied to non-linear systems of equations, the construction of higher order explicit TVD schemes remained an active area of research, see e.g. [42, 39, 17, 6] and references therein. Later, in the more general framework of strong stability preserving (SSP) implicit and explicit schemes [18], the stability property is achieved by relying on convexity arguments regarding forward and backward Euler schemes, rather than adding artificial viscosity to achieve the TVD property, as was done in [20, 42, 37]. The high-order explicit and implicit SSP schemes developed in [18, 14, 16] are limited by a very restrictive CFL condition comparable to a forward Euler scheme in order to remain oscillation-free. This makes the use of high-order implicit SSP schemes rather costly and impractical in applications, compared to high-order explicit SSP schemes, as was remarked in [16]. Regarding IMEX SSP schemes, we refer for instance to [22, 8, 23]. All high-order SSP schemes mentioned above require the time step to depend on all scales to achieve stability, but are provably high-order accurate. Unfortunately, they are not well-suited to the multi-scale setting, where the time step is strongly restricted by the fast scale leading, in extreme cases, to a vanishing time step. We also want to mention the very recent work [15], whose authors derive a high-order IMEX SSP scheme with a scale-independent time step restriction. However, this scheme cannot be applied unless some restrictive assumptions are satisfied by the system under consideration.

In contrast, our focus here is the construction of first order IMEX TVD schemes, whose CFL restriction solely stems from the explicitly treated terms associated to a scale-independent material velocity. The work presented in this manuscript is greatly motivated by the seminal work by Gottlieb et al. [18], where it was proven that an unconditionally TVD implicit RK scheme is at most first-order accurate, see also [41, 21]. Unfortunately, this result holds also for IMEX discretisations with a scale-independent CFL restriction. In fact, this discouraging result is also observed in [11, 5] when attempting to construct second-order TVD IMEX schemes for the Euler equations.

Although the TVD property is crucial to accurately capture discontinuities in the numerical solution, it becomes of less importance in smooth regions. In order to achieve a high-order approximation of the solution in such regions, while keeping the solution as oscillation-free as possible in the vicinity of discontinuities, we adapt the IMEX framework to a procedure inspired from the MOOD (Multidimensional Optimal Order Detection) techniques from [7]. The gist of the MOOD framework is to lower the order of accuracy of the scheme near problematic zones, i.e. areas where the high-order scheme violates some predetermined admissibility constraint. Therefore, a lower order scheme with good stability properties, called a *parachute scheme*, is needed in the MOOD framework.

In the present work, as stated before, we design first-order TVD IMEX RK discretisations that are consistently less diffusive than the standard first-order backward/forward Euler (IMEX1) scheme. Such discretisations therefore provide suitable fall-back schemes for the MOOD approach, yielding a reduced space-time error compared to using a traditional IMEX1 scheme as a parachute. The approach given here to construct such a fall-back scheme builds on the results from [11, 32], where the increase in precision was achieved by introducing a convex combination of said first-order TVD scheme with an oscillatory second-order scheme. In [11], the ARS(2,2,2) scheme from [2] was used as a basis for the convex combination, and this result was extended to a general class of second-order IMEX RK schemes in [32]. Here, we generalize and extend the results from [11, 32] even further, applying a convex combination to each stage, rather than merely to the final update as done in [11, 32]. This allows to construct TVD schemes based on arbitrarily high order IMEX RK schemes. Note that convex combinations have already been used to recover first-order properties lost at higher orders, see for instance [24] to recover the positivity property or [31] for well-balanced problems.

The paper is organised as follows. In Section 2, we motivate the problem of multi-scale equations, illustrated by a scalar linear transport equation. We also introduce notation for the space discretisation. Section 3 revisits the MOOD procedure and details the construction of the numerical scheme. The formalism of IMEX-RK is briefly recalled and the TVD constraints are derived for the time-semi discrete scheme. Subsequently, the fully discrete scheme is given, based on a finite volume approach and TVD limiters. To completely determine the scheme, the choice of free parameters in the convex combinations is addressed.

Section 4 is devoted to numerical experiments to verify the properties of the parachute first order TVD scheme as well as the MOOD scheme. To numerically validate that our TVD-IMEX-MOOD schemes are a noticeable improvement over widely used L-stable IMEX and the SSP IMEX schemes, we compare the performance of the schemes in terms of accuracy, CPU times and CFL restrictions on discontinuous solutions of the scalar multi-scale equation. Moreover, we numerically show that using our TVD scheme as a basis of the MOOD procedure shows a significant reduction of the space-time error. We finally apply the scheme to the isentropic Euler equations, assessing its performance for Riemann problems in two space dimensions, as well as its accuracy. To complete this manuscript, a conclusion is presented in Section 5.

2. Motivation

To study conditions to obtain a TVD scheme, we consider the linear advection equation

$$\begin{cases} w_t + c_m w_x + \dfrac{c_a}{\varepsilon} w_x = 0, \\ w(0,x) = w^0(x), \end{cases} \tag{2.1}$$

where $w : (\mathbb{R}^+, \Omega) \to \mathbb{R}$, $\Omega \subset \mathbb{R}$. In (2.1), c_m and c_a/ε denote two transport speeds which can differ significantly depending on the choice of the parameter $\varepsilon > 0$. Without loss of generality, we consider only positive transport directions, i.e. $c_m > 0$ and $c_a > 0$. In the following, we consider $c_m = \mathcal{O}(1)$ and $0 < c_a/\varepsilon \leq c_m$. The term $\frac{c_a}{\varepsilon} w_x$ is stiff as soon as $\varepsilon \ll 1$, and applying a purely explicit scheme would lead to a severe time step restriction, given by

$$\Delta t \leq \varepsilon \ \nu_{\text{ac}} \ \frac{\Delta x}{\varepsilon c_m + c_a}, \tag{2.2}$$

with a CFL coefficient ν_{ac} independent of ε. However, when ε tends to zero, the above time step vanishes, leading to huge computational costs, especially when considering long time periods. Therefore, to avoid such a restriction, we integrate the wave associated to c_a/ε implicitly, whereas $c_m w_x$ is treated explicitly. This approach leads to a CFL condition oriented to the slow wave c_m, independently of ε. It is given by

$$\Delta t \leq \nu_{\text{mat}} \ \frac{\Delta x}{c_m}. \tag{2.3}$$

In space, we use an upwind discretization since it was shown in [12] that the use of centred differences destroys the L^∞ stability for non-linear systems. Even though our considered problem is linear, the goal is to apply the scheme on non-linear systems such as the isentropic Euler equations discussed in Section 4.2.

The space and time discretisations follow the usual finite volume framework, where the computational domain Ω is divided in N uniformly spaced cells $C_j = (x_{j-1/2}, x_{j+1/2})$, of size Δx and whose center is $x_j = j\Delta x$. The solution in cell C_j is then approximated by the cell average, given by

$$w_j(t) \approx \frac{1}{\Delta x} \int_{\Omega_j} w(x,t) dx. \tag{2.4}$$

A first order space semi-discrete scheme approximating weak solutions to (2.1) is then given by

$$\partial_t w_j(t) + \frac{c_m}{\Delta x} \Delta_j(t) + \frac{c_a}{\varepsilon \Delta x} \Delta_j(t) = 0, \tag{2.5}$$

where we have introduced the abbreviation $\Delta_j(t) = w_j(t) - w_{j-1}(t)$. To obtain a fully discrete scheme, a suitable implicit-explicit time integration method has to be applied. We discretise the time variable with $t^{n+1} = t^n + \Delta t$, where Δt denotes the time step, which has to obey a material CFL condition (2.3). However, it is well-known, see e.g. [18], that approximating discontinuous solutions with high-order non-TVD methods

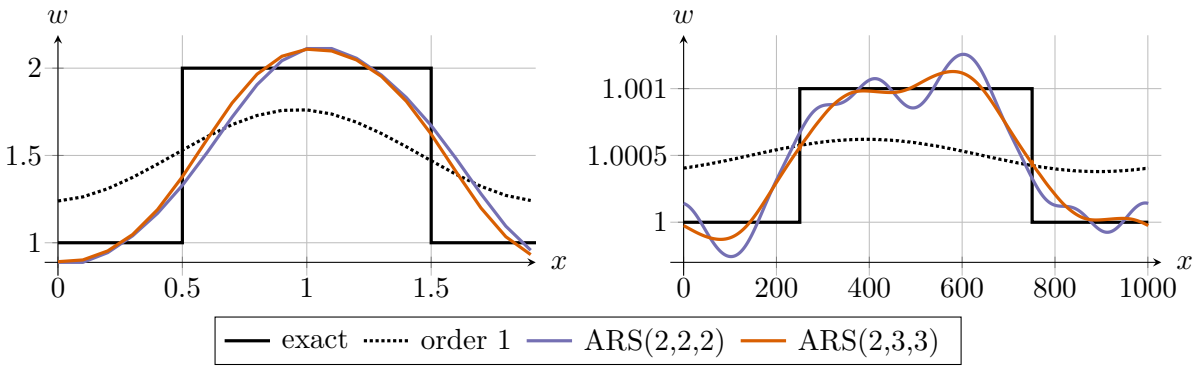


Figure 1: Approximation of a discontinuous solution with $\Delta x = 0.1$ using the first-order, second-order ARS(2,2,2) and third-order ARS(2,3,3) scheme for $\varepsilon = 1$ (left) and $\varepsilon = 10^{-3}$ (right), with an upwind space discretisation.

can lead to spurious artefacts near jump positions. This behaviour is illustrated in Figure 1, where we display the approximation of an advected rectangular bump profile with the first-order scheme

$$w_j^{n+1} = w_j^n - c_m \frac{\Delta t}{\Delta x} \Delta_j^n - \frac{c_a}{\varepsilon} \frac{\Delta t}{\Delta x} \Delta_j^{n+1}, \tag{2.6}$$

as well as with the well-known second-order ARS(2,2,2) and third-order ARS(2,3,3) IMEX schemes from [2]. The details on the numerical experiment are given in Section 4.1. We clearly observe in Figure 1 that the ARS methods violate the bounds on the numerical solution in both cases. Therefore, in order to avoid oscillations as in Figure 1, we need L^∞ stable or TVD schemes. A scheme is said to be L^∞ stable if

$$\|w^{n+1}\|_\infty = \max_{j \in \{1,\dots,N\}} |w_j^{n+1}| \leq \|w^n\|_\infty, \tag{2.7}$$

and TVD if

$$\mathrm{TV}(w^{n+1}) = \sum_{j=1}^{N} \left| w_{j+1}^{n+1} - w_j^{n+1} \right| \leq \mathrm{TV}(w^n). \tag{2.8}$$

Unfortunately, it can be proven for IMEX RK schemes, following a result of Gottlieb et al. [18], that there cannot exist L^∞ stable IMEX RK schemes of order $p \geq 2$ whose CFL restriction only stems from the explicitly treated part. Therefore, we do not look for higher order TVD IMEX integrators. However, we clearly see in Figure 1 that the first-order scheme is too diffusive to capture the structure of the solution, and thus it is also not a good choice as a base scheme in the MOOD hierarchical structure. As a consequence, our main focus here is to construct a first-order IMEX integration scheme fulfilling the L^∞ stability property (2.7) and the TVD property (2.8), that has a reduced numerical diffusion compared to the first order scheme (2.6), and is therefore suited as a fallback scheme when combined with a MOOD procedure.

3. Derivation of the numerical scheme

The goal of this section is the derivation of a numerical scheme based on a MOOD-like procedure. The usual MOOD framework for explicit schemes, see e.g. [7], consists in locally and gradually lowering the order of the scheme when an oscillation is detected. The lowest order scheme the procedure can use is called a *parachute scheme*. Generally speaking, it guarantees the preservation of desired properties not satisfied by the higher order schemes. Here, the higher order scheme is given by a standard IMEX-RK scheme, and the lowest order scheme consists of a first order TVD-IMEX scheme. Consequently, the precision of the first order TVD scheme can be improved without degrading its stability properties, yielding a high order approximation in smooth flow regimes. Due to the non-local nature of the implicit part in the IMEX schemes,

the standard MOOD algorithm as given in [7] has to be modified, since the solution cannot be recomputed on a few selected cells only. Therefore, it has to be updated on the whole mesh.

Depending on the application, different detection criteria can be applied to identify oscillations violating the TVD property. For the toy problem (2.1), an oscillation is detected when the solution leaves the bounds of the initial condition. Indeed, the unknown w for this toy problem satisfies the maximum principle $\partial_t \|w(t,x)\|_\infty \leq 0$, and therefore an L^∞-stable discretisation obeys $\|w^{n+1}\|_\infty \leq \|w^0\|_\infty$ for all $n \geq 0$. Thus, the detection criterion, denoted by Φ, depends on the initial condition only and is time independent. However, in general, time dependent detection criteria must be taken into account for non-linear problems, as detailed for the isentropic Euler equations in Section 4. The modified MOOD scheme, see also [11, 32], is given by the following algorithm.

Algorithm 1 (MOODp scheme)**.** *Define the initial detection criterion given by $\mathcal{E}^0 = \|\Phi(w^0)\|$. Equipped with a stable first order TVD scheme, the MOODp scheme consists in applying the following procedure at each time step:*

(1) Compute a candidate numerical solution w_c^{n+1} with a p^{th} order IMEX-RK scheme.

(2) Detect whether an oscillation is present somewhere in the space domain, i.e. whether the detection criterion is satisfied by the candidate solution:

$$\|\Phi(w_c^{n+1})\|_\infty \leq \mathcal{E}^n. \tag{DMP}$$

(3a) If (DMP) holds, set for the numerical solution at the new time step $w^{n+1} = w_c^{n+1}$.

(3b) Otherwise, compute a solution $w_{\Parachute}^{n+1}$ with the parachute scheme and set $w^{n+1} = w_{\Parachute}^{n+1}$.

(4) For a time-dependent detection criterion, update $\mathcal{E}^{n+1} = \xi\|\Phi(w^{n+1})\|_\infty + (1-\xi)\mathcal{E}^n$ with $\xi \in [0,1]$, otherwise set $\mathcal{E}^{n+1} = \mathcal{E}^0$.

Note that, in Step (4) of the MOOD algorithm, the detection criterion is relaxed with a convex combination of parameter $\xi \in [0,1]$ between the current solution and the criterion at the previous time step. This allows a finer control of the permitted oscillations in the MOOD solution.

A crucial part of the MOOD scheme is the construction of the parachute scheme, which should preserve the TVD property of the solution in case the higher order scheme produces oscillations. This is addressed in the following section.

3.1. Construction of the time-semi discrete parachute scheme

The proposed parachute schemes are based on an IMEX-RK approach where the implicit part is diagonally implicit. The time update for an s-stage IMEX-RK scheme for equation (2.5) is given by

$$w_j^{n+1} = w_j^n - \lambda \sum_{k=1}^{s} \tilde{b}_k \Delta_j^{(k)} - \mu_\varepsilon \sum_{k=1}^{s} b_k \Delta_j^{(k)}, \tag{3.1}$$

where we have set

$$\lambda = \frac{\Delta t}{\Delta x} c_m, \qquad \mu_\varepsilon = \frac{\Delta t}{\Delta x}\frac{c_a}{\varepsilon}, \qquad \Delta_j^{(k)} = w_j^{(k)} - w_{j-1}^{(k)},$$

and where the stages are defined as

$$w_j^{(k)} = w_j^n - \lambda \sum_{l=1}^{k-1} \tilde{a}_{kl} \Delta_j^{(l)} - \mu_\varepsilon \sum_{l=1}^{k} a_{kl} \Delta_j^{(l)}. \tag{3.2}$$

For the sake of clarity, we consider an IMEX-RK method of type CK (Carpenter and Kennedy) [25], i.e. we take $a_{11} = 0$. For results based on a non-CK scheme with $a_{11} \neq 0$ see Appendix Appendix A. In order to obtain a CFL condition like (2.3) which does not depend on ε, the weights $(a_{k1})_{k\in\{2,\dots,s\}}$, as well as b_1,

have to be zero. This was shown in detail in [32] for a generic second-order CK method. We summarize the structure of the RK scheme in the following Butcher tableaux notation

$$\text{explicit:}\quad \begin{array}{c|cccc} 0 & 0 & 0 & \cdots & 0 \\ \tilde{c}_2 & \tilde{a}_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \tilde{c}_s & \tilde{a}_{s1} & \cdots & \tilde{a}_{s,s-1} & 0 \\ \hline & \tilde{b}_1 & \cdots & \tilde{b}_{s-1} & \tilde{b}_s \end{array} \qquad \text{implicit:}\quad \begin{array}{c|cccc} 0 & 0 & 0 & \cdots & 0 \\ c_2 & 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & 0 & a_{s2} & \cdots & a_{ss} \\ \hline & 0 & b_2 & \cdots & b_s \end{array}, \tag{3.3}$$

with the coefficients $\tilde{c}$ and c obeying

$$\tilde{c}_i = \sum_{j=1}^{i-1} \tilde{a}_{ij} \quad \text{and} \quad c_i = \sum_{j=1}^{i} a_{ij}. \tag{3.4}$$

Since the TVD scheme we construct are based on higher order IMEX-RK schemes, the weights have to fulfil high-order compatibility conditions as given in [34]. For orders higher than three, we refer to the order conditions in [25].

To obtain a TVD scheme from an s-stage IMEX-RK scheme, we propose a convex combination of each stage with the first-order IMEX Euler scheme (2.6) at time $t^n + c_k\Delta t$ for the k^{th} stage. This approach yields s free parameters $\theta_k \in [0,1]$, where $k = 1,\dots,s$ denotes the stage in the IMEX scheme. The closer each θ_k is to 1, the higher the contribution of the IMEX-RK scheme. As a consequence, we shall seek to maximize the values of θ_k, to reduce diffusion as much as possible compared to the first-order IMEX Euler scheme. The stages of the TVD scheme are defined by

$$w_j^{(k)} + (1-\theta_k)c_k\mu_\varepsilon\Delta_j^{(k)} = w_j^n - \lambda\left((1-\theta_k)\tilde{c}_k\Delta_j^n + \theta_k\sum_{l=1}^{k-1}\tilde{a}_{kl}\Delta_j^{(l)}\right) - \mu_\varepsilon\theta_k\sum_{l=1}^{k}a_{kl}\Delta_j^{(l)}, \tag{3.5}$$

and the update by

$$w_j^{n+1} = w_j^n - \theta_{s+1}\left(\lambda\sum_{k=1}^{s}\tilde{b}_k\Delta_j^{(k)} + \mu_\varepsilon\sum_{k=1}^{s}b_k\Delta_j^{(k)}\right) - (1-\theta_{s+1})\left(\lambda\Delta_j^n + \mu_\varepsilon\Delta_j^{n+1}\right). \tag{3.6}$$

Note that, for Butcher tableaux like (3.3), we immediately set $\theta_1 = 1$ to recover $w^{(1)} = w^n$. This means that the convex stages appear earliest for $k = 2$. In the case where the weights $\tilde{b}$ and b respectively coincide with the last row of $\tilde{A}$ and A, the stage $w^{(s)}$ coincides with the final update w^{n+1}. In particular, we then have $\theta_{s+1} = \theta_s$.

We emphasise that a TVD scheme resulting from the convex combination (3.6) is at most first-order accurate, see for instance [41, 18, 21]. However, it is a great fit as a parachute scheme for the MOOD Algorithm 1. Section 3.1.1 is dedicated to the necessary conditions on the free parameters in order to obtain the TVD property using a three stage IMEX-RK scheme. Then, in Section 3.1.2, we state an extension to construct TVD discretisations based on more general Butcher tableaux (3.3).

3.1.1. TVD conditions for a three stage scheme

As basis of the TVD scheme, we consider Butcher tableaux with $s_{\text{eff}} = 3$ effective computational steps. In the spirit of (3.3), they are given by

$$\text{explicit:}\quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ c_2 & \tilde{a}_{21} & 0 & 0 \\ c_3 & \tilde{a}_{31} & \tilde{a}_{32} & 0 \\ \hline & 0 & b_2 & b_3 \end{array}, \qquad \text{implicit:}\quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ c_2 & 0 & a_{22} & 0 \\ c_3 & 0 & a_{32} & a_{33} \\ \hline & 0 & b_2 & b_3 \end{array}. \tag{3.7}$$

Therein, we have assumed that $\tilde{b} = b$ and $\tilde{c} = c$, see also [35]. This has the advantage that the weights in the final update coincide with respect to the explicitly and implicitly treated terms. Applying the third-order conditions as given in [34] on the weights and coefficients given by (3.7) lead to the following tableaux, with $\gamma \notin \{0, \frac{1}{3}\}$:

$$\text{explicit:} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{3\gamma-1}{6\gamma} & \frac{3\gamma-1}{6\gamma} & 0 & 0 \\ \frac{\gamma+1}{2} & -\frac{6\gamma^3-3\gamma^2+1}{2(3\gamma-1)} & \frac{\gamma(3\gamma^2+1)}{3\gamma-1} & 0 \\ \hline & 0 & \frac{3\gamma^2}{3\gamma^2+1} & \frac{1}{3\gamma^2+1} \end{array} \qquad \text{implicit:} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{3\gamma-1}{6\gamma} & 0 & \frac{3\gamma-1}{6\gamma} & 0 \\ \frac{\gamma+1}{2} & 0 & \gamma & \frac{1-\gamma}{2} \\ \hline & 0 & \frac{3\gamma^2}{3\gamma^2+1} & \frac{1}{3\gamma^2+1} \end{array} \tag{3.8}$$

We now demonstrate that, using Butcher tableaux (3.8), a first order TVD scheme can be obtained. First, we show the L^∞ stability (2.7), i.e. $\|w^{n+1}\|_\infty \leq \|w^n\|_\infty$. The idea of the proof lies in estimating the L^∞ norm of each stage against $\|w^n\|_\infty$, which will then be used to obtain the final estimate (2.7). The proof is achieved by only applying the triangle and reverse triangle inequalities. For clarity, we do not replace the coefficients of the tableaux by their γ-dependent values yet.

The first stage reduces to $w^{(1)} = w^n$ and we trivially obtain $\|w^{(1)}\|_\infty = \|w^n\|_\infty$. Furthermore, $w^{(1)}$ is independent of θ_1, and therefore we set $\theta_1 = 1$, as mentioned before. Using definition (3.5), the second stage is given by

$$w_j^{(2)} + a_{22}\mu_\varepsilon \Delta_j^{(2)} = w_j^n - \lambda \left((1-\theta_2) a_{22} \Delta_j^n + \theta_2 \tilde{a}_{21} \Delta_j^n \right).$$

Reformulating and collecting terms in w_j^n, w_{j-1}^n, and using that $\tilde{a}_{21} = a_{22}$, we obtain the following expression

$$w_j^{(2)} + a_{22}\mu_\varepsilon \Delta_j^{(2)} = (1 - \lambda a_{22})\, w_j^n + \lambda a_{22} w_{j-1}^n. \tag{3.9}$$

Note that the second stage, like the first one, is independent of θ_2. Thus, we set $\theta_2 = 1$, which achieves the full contribution of the second stage of the IMEX-RK scheme. For periodic boundary conditions and requiring a_{22} and $1 - \lambda a_{22}$ to be positive, we obtain the following estimates

$$\begin{aligned} \|w^n\|_\infty &= (1-\lambda a_{22})\, \|w^n\|_\infty + \lambda a_{22} \|w^n\|_\infty = (1 - \lambda a_{22}) \max_j |w_j^n| + \lambda a_{22} \max_j |w_{j-1}^n| \\ &\geq \max_j \left| (1-\lambda a_{22})\, w_j^n + \lambda a_{22} w_{j-1}^n \right| = \max_j \left| w_j^n - \lambda a_{22} (w_j^n - w_{j-1}^n) \right| \\ &= \max_j \left| (1 + \mu_\varepsilon a_{22}) w_j^{(2)} - \mu_\varepsilon a_{22} w_{j-1}^{(2)} \right| \geq (1 + \mu_\varepsilon a_{22}) \|w^{(2)}\|_\infty - \mu_\varepsilon a_{22} \|w^{(2)}\|_\infty \\ &= \|w^{(2)}\|_\infty. \end{aligned} \tag{3.10}$$

Further, from $a_{22} > 0$ we immediately obtain a constraint on γ given by $\frac{3\gamma-1}{6\gamma} > 0$, i.e. $\gamma > \frac{1}{3}$ or $\gamma < 0$. Moreover, from $1 - \lambda a_{22} > 0$ we obtain a restriction on the time step given by $\lambda < \frac{6\gamma}{3\gamma-1}$.

Turning to the third stage of the scheme given in (3.5), we obtain

$$\begin{aligned} w_j^{(3)} + (1-\theta_3) c_3 \mu_\varepsilon \Delta_j^{(3)} = w_j^n &- \lambda \left((1-\theta_3) \tilde{c}_3 \Delta_j^n + \theta_3 \left(\tilde{a}_{31} \Delta_j^{(1)} + \tilde{a}_{32} \Delta_j^{(2)} \right) \right) \\ &- \mu_\varepsilon \theta_3 \left(a_{32} \Delta_j^{(2)} + a_{33} \Delta_j^{(3)} \right). \end{aligned} \tag{3.11}$$

The next step consists in eliminating the terms in $\mu_\varepsilon w^{(2)}$ to avoid a ε restriction on λ. To that end, we invoke (3.9) to replace $-\mu_\varepsilon \Delta_j^{(2)}$ in (3.11) by

$$-\mu_\varepsilon \Delta_j^{(2)} = \frac{1}{a_{22}} (w_j^{(2)} - w_j^n) + \lambda \Delta_j^n, \tag{3.12}$$

which yields, after some rearranging,

$$\begin{aligned} w_j^{(3)} + \mu_\varepsilon \left((1-\theta_3) c_3 + \theta_3 a_{33} \right) \Delta_j^{(3)} = w_j^n &- \lambda \left((1-\theta_3) c_3 \Delta_j^n + \theta_3 \left(\tilde{a}_{31} \Delta_j^n + \tilde{a}_{32} \Delta^{(2)} \right) \right) \\ &+ \theta_3 a_{32} \left(\frac{1}{a_{22}} (w_j^{(2)} - w_j^n) + \lambda \Delta_j^n \right). \end{aligned} \tag{3.13}$$

Reformulating (3.13) to highlight the contribution of each step gives

$$\begin{aligned} w_j^{(3)} + \mu_\varepsilon \left(c_3 + \theta_3(a_{33} - c_3)\right) \Delta_j^{(3)} = &\left(1 - \lambda(1-\theta_3)c_3 - \theta_3\lambda(\tilde{a}_{31} - a_{32}) - \frac{\theta_3 a_{32}}{a_{22}}\right) w_j^n \\ &+ \left(\lambda(1-\theta_3)c_3 + \theta_3\lambda(\tilde{a}_{31} - a_{32})\right) w_{j-1}^n \\ &+ \theta_3 \left(\frac{a_{32}}{a_{22}} - \lambda\tilde{a}_{32}\right) w_j^{(2)} + \theta_3\ \lambda\tilde{a}_{32} w_{j-1}^{(2)}. \end{aligned}$$

To obtain the estimate $\|w^n\|_\infty \geq \|w^{(3)}\|_\infty$, we proceed analogously to the proof for the second stage. The proof follows the lines of (3.10) and is thus omitted. Like in the second stage, we find that the coefficients in front of all stages involving w_j and w_{j-1} have to be non-negative. The non-negativity requirements on the coefficients of $w_j^{(2)}$ and $w_{j-1}^{(2)}$ result into the conditions

$$\begin{aligned} \tilde{a}_{32} \geq 0 \iff \frac{\gamma(3\gamma^2+1)}{3\gamma - 1} \geq 0 \iff \gamma \leq 0 \text{ or } \gamma > \frac{1}{3}, \\ \frac{a_{32}}{a_{22}} - \lambda\tilde{a}_{32} \geq 0 \iff \lambda \leq \frac{6\gamma^2}{\gamma(3\gamma^2+1)} \text{ and } \gamma > \frac{1}{3}. \end{aligned} \tag{3.14}$$

Further we find, thanks to the coefficient in front of w_{j-1}^n,

$$(1-\theta_3)c_3 + \theta_3(\tilde{a_{31}} - a_{32}) \geq 0 \iff \theta_3 \frac{3\gamma^2(\gamma+1)}{3\gamma-1} \leq \frac{\gamma+1}{2} \iff \theta_3 \leq \frac{3\gamma-1}{6\gamma^2}, \tag{3.15}$$

which yields a restriction on θ_3 depending on γ. Another estimate for θ_3 can be obtained from the coefficient of $\Delta_j^{(3)}$ given by

$$c_3 + \theta_3(a_{33} - c_3) \geq 0 \iff \gamma\theta_3 \leq \frac{\gamma+1}{2} \iff \theta_3 \leq \frac{\gamma+1}{2\gamma}.$$

We see that this condition on θ_3 is less restrictive than the one obtained from (3.15) for all $\gamma > \frac{1}{3}$. Note that $\gamma < 0$ would yield a negative θ_3, which imposes $\gamma > 0$. The largest value we can take for θ_3 is therefore given by

$$\theta_3^{\text{opt}} = \frac{3\gamma - 1}{6\gamma^2},$$

and θ_3 must satisfy $\theta_3 \leq \theta_3^{\text{opt}}$. Note that the coefficient of w_j^n is always non-negative. The expression

$$1 - \lambda(1-\theta_3)c_3 - \theta_3\lambda(\tilde{a}_{31} - a_{32}) - \frac{\theta_3 a_{32}}{c_2} \geq 0 \tag{3.16}$$

is always fulfilled if we set $\theta_3 = \theta_3^{\text{opt}}$. Using θ_3^{opt} yields the maximal allowed input from the stages (3.2) of the third order IMEX-RK scheme. Otherwise, (3.16) leads to another, more restrictive estimate for λ. Having obtained the estimate for stage three, we turn to the final update given by

$$w_j^{n+1} + (1-\theta_4)\mu_\varepsilon \Delta_j^{n+1} = w_j^n - \theta_4(\lambda + \mu_\varepsilon) b_2 \Delta_j^{(2)} - \theta_4(\lambda+\mu_\varepsilon) b_3 \Delta_j^{(3)} - (1-\theta_4)\lambda\Delta_j^n. \tag{3.17}$$

We again eliminate the terms in $\mu_\varepsilon\Delta_j^{(2)}$ and $\mu_\varepsilon\Delta_j^{(3)}$ in order to avoid a CFL restriction depending on ε. In (3.17), we replace $-\mu_\varepsilon\Delta_j^{(2)}$ by (3.12) and $-\mu_\varepsilon\Delta_j^{(3)}$, using (3.13), by

$$\begin{aligned} -\mu_\varepsilon\Delta_j^{(3)} = \frac{1}{c_3 + \theta_3(a_{33}-c_3)} \left(w_j^{(3)} - w_j^n\right) + \frac{c_3}{c_3 + \theta_3(a_{33}-c_3)} \lambda \left(\left(1 + \theta_3(\tilde{a}_{31} - 1)\right) \Delta_j^n + \theta_3\tilde{a}_{32}\Delta_j^{(2)}\right) \\ - \frac{\theta_3 c_3 a_{32}}{c_3 + \theta_3(a_{33}-c_3)} \left(\frac{1}{a_{22}}(w_j^{(2)} - w_j^n) + \lambda\Delta_j^n\right). \end{aligned}$$

Collecting the terms in front of the states w_j, w_{j-1} at each stage and requiring the positivity of these coefficients gives the L^∞ property $\|w^{n+1}\|_\infty \leq \|w^n\|_\infty$ following analogous steps to (3.10). The resulting requirements on the CFL condition λ and the convex parameter θ_4 are given in Lemma 1.

The TVD property can be obtained following the same line of computations as performed for the L^∞ stability and is contained in the following result.

Lemma 1 (L^∞ stability, TVD property)**.** *For periodic boundary conditions under the CFL condition*

$$\lambda \leq \frac{18\gamma^3\theta_4 - (3\gamma-1)(3\gamma^2+1)}{(3\gamma-1)((6\gamma^2+1)\theta_4 - (3\gamma^2+1))},$$

the scheme consisting of the Butcher tableaux (3.8) with the convex scheme given by the stages (3.6) and the update (3.5) is L^∞ stable and TVD if the following conditions are fulfilled:

$$\gamma \geq \frac{\sqrt{3}}{3}, \quad \theta_1 = 1, \quad \theta_2 = 1, \quad \theta_3 = \frac{3\gamma-1}{6\gamma^2}, \quad \theta_4 < \frac{(3\gamma-1)(3\gamma^2+1)}{18\gamma^3}. \tag{3.18}$$

3.1.2. TVD conditions for arbitrary number of stages

In the spirit of the results from the third-order scheme, we now seek a general framework on how to obtain TVD schemes with s stages using the IMEX formulation (3.5)–(3.6). Since the proof follows steps analogous to the ones from Section 3.1.1, we do not repeat the calculations and directly give the final result.

Theorem 2. *Let $\tilde{A}, A \in \mathbb{R}^{s\times s}$, $\tilde{b}, b, \tilde{c}, c \in \mathbb{R}^s$ define two Butcher tableaux (3.3) fulfilling (3.4) and the p-th order compatibility conditions. Let $\tilde{b}$ and b coincide with the last rows of $\tilde{A}$ and A respectively. For $k = 1, \dots, s$ and $l = 1, \dots, k-1$, we define*

$$\mathcal{A}_k = \theta_k a_{kk} + (1-\theta_k)c_k, \quad \tilde{\mathcal{A}}_k = \theta_k a_{k1} + (1-\theta_k)\tilde{c}_k, \quad \mathcal{B}_{kl} = \frac{\theta_k a_{kl}}{\mathcal{A}_l}, \quad \tilde{\mathcal{B}}_{kl} = \theta_k \tilde{a}_{kl}.$$

In addition, we recursively define the following expressions:

$$\mathcal{C}_k = \tilde{\mathcal{A}}_k - \sum_{l=2}^{k-1} \mathcal{B}_{kl}\mathcal{C}_l, \qquad \mathcal{C}_{kl} = \tilde{\mathcal{B}}_{kl} - \sum_{r=l+1}^{k-1} \mathcal{B}_{kr}\mathcal{C}_{rl},$$

$$\mathcal{D}_k = 1 - \lambda\tilde{\mathcal{A}}_k - \sum_{l=2}^{k-1} \mathcal{B}_{kl}\mathcal{D}_l, \qquad \mathcal{D}_{kl} = \mathcal{B}_{kl} - \lambda\tilde{\mathcal{B}}_{kl} - \sum_{r=l+1}^{k-1} \mathcal{B}_{kr}\mathcal{D}_{rl}.$$

Then, with $\theta_1 = 1$ and under the following restrictions for $k = 2, \dots, s$ and $l = 1, \dots, k-1$,

$$\mathcal{A}_k > 0, \quad \mathcal{C}_k \geq 0, \quad \mathcal{D}_k \geq 0, \quad \mathcal{C}_{kl} \geq 0, \quad \mathcal{D}_{kl} \geq 0.$$

the scheme consisting of the stages (3.5) and the update (3.6), combined with a TVD limiter, is L^∞ stable and TVD under a CFL condition determined by $\lambda \geq 0$ where λ does not depend on ε.

The result from Theorem 2 can be extended to the case where the weights $\tilde{b}$ and b do not coincide with the respective last rows of $\tilde{A}$ and A. To be able to use the notation from Theorem 2, we view the update (3.6) as an additional explicit $(s+1)$-th stage of a scheme induced by Butcher tableaux (3.3), with $(s+1)\times(s+1)$ matrices with the diagonal entry $a_{s+1,s+1} = 0$, and where the weights $\tilde{b}$ and b respectively coincide with the last rows of the new $\tilde{A}$ and A. Theorem 2 is then applied to yield the L^∞ stability and the TVD property. We conclude this section with some remarks.

Remark 1. We can prove the same result if the first column of A allows for non-zero entries. The TVD conditions obtained when assuming that structure are given in Appendix Appendix A.

Remark 2. Applying this procedure to a two-stage, second-order scheme, we recover the ARS(2,2,2)-based schemes introduced in [11] and studied more broadly in [32].

3.2. Fully discrete parachute scheme

To increase the resolution of the explicit spatial derivatives, we provide a third-order reconstruction of the cell interface values $w_{j+1/2}$ such that the fully discrete scheme is L^∞ stable and TVD. We reconstruct the values $w_j^{(k)}$ using the neighbouring cell averages, see for instance [28]. The reconstructed values $w_{j,-}^{(k)}$ and $w_{j,+}^{(k)}$ at the inner interfaces of cell C_j are then defined by

$$w_{j,-}^{(k)} = w_j^{(k)} - \frac{\Delta x}{2} L\left(\sigma_{j+1/2}^{(k)}, \sigma_{j-1/2}^{(k)}\right), \quad w_{j,+}^{(k)} = w_j^{(k)} + \frac{\Delta x}{2} L\left(\sigma_{j-1/2}^{(k)}, \sigma_{j+1/2}^{(k)}\right), \tag{3.19}$$

where $\sigma_{j+1/2}^{(k)}$ denotes the slope between the values of $w_j^{(k)}$ and $w_{j+1}^{(k)}$ given by

$$\sigma_{j+1/2}^{(k)} = \frac{w_{j+1}^{(k)} - w_j^{(k)}}{\Delta x}.$$

The function $L(\sigma_L, \sigma_R)$ is a slope limiter which ensures that the reconstructed values still satisfy L^∞ property. For a three-point stencil, the following estimate has to hold

$$\min(|w_{j-1}^{(k)}|, |w_j^{(k)}|, |w_{j+1}^{(k)}|) \leq |w_{j,\pm}^{(k)}| \leq \max(|w_{j-1}^{(k)}|, |w_j^{(k)}|, |w_{j+1}^{(k)}|). \tag{3.20}$$

Here, we use a TVD third-order space reconstruction satisfying (3.20) by following the limiting procedure introduced in [38]. This procedure switches between the oscillatory unlimited third-order reconstruction and a third-order TVD limiter. Switching to the TVD limiter is triggered by the appearance of a non-physical oscillation represented by a non-smooth extremum.

In the spirit of the reconstruction used to approximate the explicit derivatives, we could also increase the space accuracy of the implicit derivatives using TVD slope limiters. Note that the slopes are determined in general by a non-linear function. This implies an implicit approximation of the reconstructed values (3.19). Such computations usually lead to an iterative process or a prediction-correction method, and are therefore extremely costly. We consider this increase in computational cost as too much in the sight of the actual gain in resolution. Note that the IMEX-TVD scheme is overall only of first order. Therefore, this is a loss of resolution in space we are willing to take to obtain a less costly TVD scheme.

3.3. Choice of the free parameters

The TVD scheme of Section 3.1.1, denoted by TVD3, contains some free parameters. We now suggest optimal values of these free parameters. To that end, we analyse the error produced by the schemes, as well as the CPU time taken, with respect to the free parameters. This analysis will help us give some insights on how to optimally choose these parameters, and on the trade-offs that must be made when making such choices.

We compare the IMEX1 scheme to the IMEX3, TVD3 and MOOD3 schemes. Here, the IMEXp scheme is the unlimited scheme of order p in space and time. Following this notation, the IMEX1 scheme is given by (2.6) and the IMEX3 scheme corresponds to the Butcher tableaux (3.8). The MOODp scheme is obtained in the framework of Algorithm 1, using the IMEXp scheme as the high-order scheme and the TVDp scheme as the parachute scheme. For the linear multi-scale advection equation, the detection criterion consists in taking the L^∞ norm of the solution, and thus we take $\Phi(w) = w$ in Algorithm 1. We also take $\xi = 0$ to eliminate all oscillations.

Since the scope of this section is to study the effect of the time discretisation on the precision and computational time of our schemes, we temporarily restrict ourselves to a first-order upwind space discretisation. This ensures that only the effects of the time discretisation are monitored.

An exact smooth solution of the toy problem (2.1) is used for the calibration of the free parameters $\gamma, \theta_3, \theta_4$ and λ in the TVD3 scheme. It is given by

$$w^s(t,x) = 1 + \frac{\varepsilon}{2}\left(1 + \sin\left[2\pi\varepsilon\left(x - \left(c_m + \frac{c_a}{\varepsilon}\right)t\right)\right]\right), \tag{3.21}$$

which represents a sine function of amplitude ε, transported with the velocity $c_m + \frac{c_a}{\varepsilon}$. We set $\varepsilon = 0.1$ and we take $N = 400$ cells. The conclusions of the forthcoming developments are unchanged if we consider other values of ε. Indeed, taking a different ε would merely translate the curves without changing their relative positions.

3.3.1. Choice of γ and θ_4 in the TVD3 scheme

For the TVD3 scheme, we have to set the values of θ_3, θ_4 and λ, constrained by Lemma 1. Ideally, we would like θ_3, θ_4 and λ to be as large as possible. By inspection, we note that the maximum value of θ_3 is $\theta_3^{\text{opt}} = \frac{3}{8}$, obtained for $\gamma^{\text{opt}} = \frac{2}{3}$. The Butcher tableaux (3.8) then become

$$\text{explicit:} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 0 \\ 5/6 & -13/18 & 14/9 & 0 \\ \hline & 0 & 4/7 & 3/7 \end{array}, \qquad \text{implicit:} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/4 & 0 & 1/4 & 0 \\ 5/6 & 0 & 2/3 & 1/6 \\ \hline & 0 & 4/7 & 3/7 \end{array}. \tag{3.22}$$

Taking this value of γ in Lemma 1 yields the following bounds on θ_4 and λ

$$0 < \theta_4 < \frac{7}{16} \quad \text{and} \quad 0 < \lambda < \frac{7 - 16\theta_4}{7 - 11\theta_4}. \tag{3.23}$$

We note that λ is a decreasing function of θ_4, which implies that we are not able to use both a large θ_4 and a large λ. There is a trade-off between the CFL condition λ, i.e. the CPU time, and the value of θ, i.e. the resolution of the scheme. To quantify this balance between precision and CPU time, let us introduce $\alpha \in (0, 1)$, to rewrite (3.23) as follows

$$\theta_4 = \frac{7}{16}\alpha \quad \text{and} \quad \lambda = \frac{1 - \alpha}{1 - \frac{11}{16}\alpha}. \tag{3.24}$$

We note that θ_4 increases and λ decreases with increasing α. Making use of the formulation in terms of α, we analyse the TVD3 scheme with $\gamma = \gamma^{\text{opt}} = \frac{2}{3}$. Note that a similar study was performed, for the second-order case, in [32]. We first display in Figure 2 the CPU time with respect to α for the four schemes. As expected, since the CFL condition becomes more restrictive, the CPU time increases with α for the TVD3 and the MOOD3 schemes. Second, in the left panel of Figure 3, we display the L^∞-error with respect to α for the four schemes under consideration. As expected, we observe that it decreases with α for the TVD3 scheme, since θ_4 increases. Third, in the right panel of Figure 3, we display a zoom on the CPU time and the

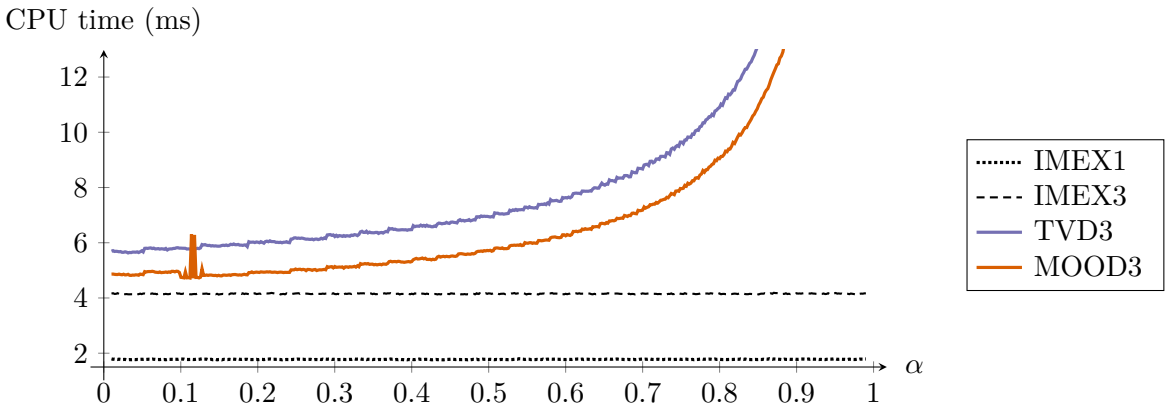


Figure 2: CPU time (in milliseconds) with respect to the parameter α, using $\gamma = \gamma^{\text{opt}} = \frac{2}{3}$, in the context of the test case presented in Section 3.3.1.

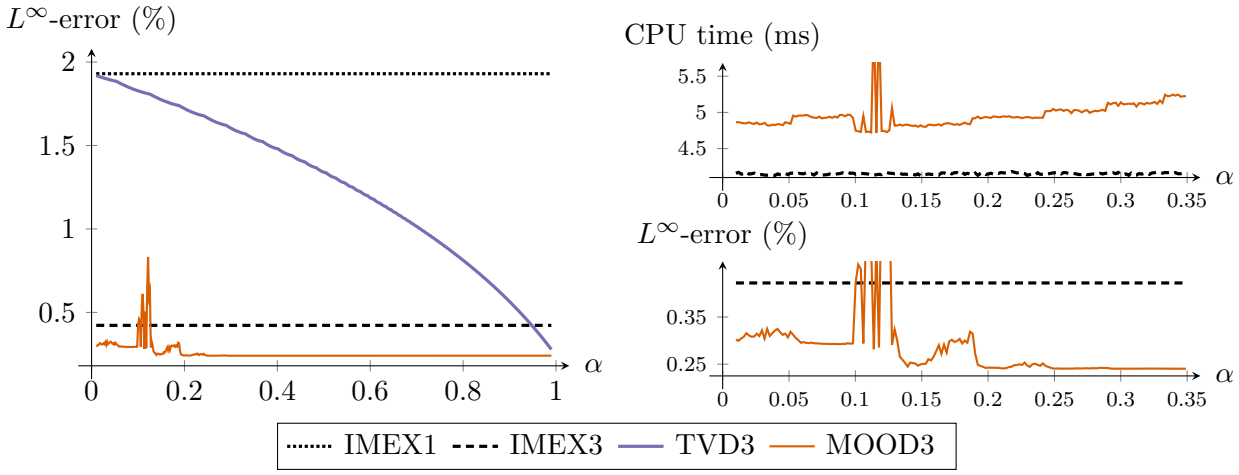


Figure 3: L^∞-error with respect to the parameter α, using $\gamma = \gamma^{\text{opt}} = \frac{2}{3}$, $\theta_3 = \frac{3}{8}$ and θ_4, λ given by relation (3.24). in the context of the test case presented in Section 3.3.1. For $\alpha \in (0, 0.35)$, the top right panel contains a zoom on the CPU time (data from Figure 2) and the bottom right panel contains a zoom on the L^∞-error (data from left panel).

L^∞-error produced by the IMEX3 and MOOD3 schemes, with respect to $0 < \alpha < 0.35$. We observe that the error stabilizes around $\alpha = 0.3$, and that the CPU time increases monotonically with α. Therefore, taking $\alpha = \frac{1}{3}$ seems to be a good compromise between precision and computational time. In the remainder of this article, we take

$$\gamma = \gamma^{\text{opt}} = \frac{2}{3} \quad \text{and} \quad \alpha = \alpha^{\text{opt}} = \frac{1}{3},$$

which leads to the following values for θ_3, θ_4 and λ

$$\theta_3^{\text{opt}} = \frac{3}{8} = 0.375, \quad \theta_4^{\text{opt}} = \frac{7}{48} \simeq 0.146, \quad \text{and} \quad \lambda^{\text{opt}} = \frac{32}{37} \simeq 0.865.$$

3.3.2. Numerical optimisation of larger Butcher tableaux

To conclude this section, we mention a four-step third-order Butcher tableau yielding a TVD scheme. To obtain this tableau, we have used the TVD inequalities from Theorem 2, as well as the order conditions as constraints in an optimisation problem. Its objective is to maximize the value of $\lambda + \sum \theta$, and its unknowns are the Butcher coefficients, the values of θ, and λ. We ran this optimization problem with many random initial conditions for the unknowns, and we refined this random initialisation around values yielding a large value of the objective function. In the end, we chose the solution where the value of the objective function was maximal, under the additional constraint that $\lambda \geq 0.5$, which is a standard CFL condition arising in fluid dynamics schemes. We obtained $\lambda = 0.5471076190680170$, $\theta_1 = 1$, $\theta_2 = 1$,$\theta_3 = 1$, $\theta_4 = 0.5110907014643069$ and $\theta_5 = 0.4997722865197203$. The Butcher tableau is given in Appendix Appendix B. In the remainder of the paper, the scheme and its MOOD version will be referred to as TVD3(4) and MOOD3(4).

4. Numerical results

In this section, we verify the properties of the numerical schemes that were developed in the previous sections. The schemes are summarized in Table 1. We first apply the TVD-MOOD strategy on the scalar linear problem from Section 4.1, before moving on to nonlinear systems of equations in Section 4.2. We consider the 2D isentropic Euler equations as an example of such a system. We recall the two different types of CFL conditions used in the following. The first one is given by an acoustic CFL condition

$$\Delta t \leq \nu_{\text{ac}} \frac{\Delta x}{\max_k |\lambda_k|}, \tag{4.1}$$

First-order IMEX (2.6)	Scheme from Section 3.3.1		Scheme from Section 3.3.2	
parachute: IMEX1	unlimited: parachute: MOOD version:	IMEX3 TVD3 MOOD3	unlimited: parachute: MOOD version:	IMEX3(4) TVD3(4) MOOD3(4)

Table 1: Names of the schemes used in the numerical simulations.

which is restricted by the fastest wave speed $\max|\lambda_k|$. The second one is a material CFL condition

$$\Delta t \leq \nu_{\text{mat}} \frac{\Delta x}{|\lambda_u|}, \tag{4.2}$$

which is only constrained by the system velocity $|\lambda_u| \leq \max|\lambda_k|$ and thus allows larger time steps than the acoustic CFL condition. For the toy problem, $\max|\lambda_k| = c_m + \frac{c_a}{\varepsilon}$ and $\lambda_u = c_m$ according to (2.2) and (2.3).

4.1. Linear multi-scale advection equation

Since oscillations typically appear around discontinuities, we consider the following exact solution of the initial value problem (2.1)

$$w(t,x) = \begin{cases} 1+\varepsilon & \text{if } \dfrac{1}{4} < \left(\dfrac{\left(x - \left(c_m + \frac{c_a}{\varepsilon}\right) t\right)}{c_m + \frac{c_a}{\varepsilon}} - \left\lfloor \dfrac{\left(x - \left(c_m + \frac{c_a}{\varepsilon}\right) t\right)}{c_m + \frac{c_a}{\varepsilon}} \right\rfloor \right) < \dfrac{3}{4}, \\ 1 & \text{otherwise,} \end{cases} \tag{4.3}$$

which represents a rectangular bump located at $\left(\frac{1}{4}(c_m + \frac{c_a}{\varepsilon}), \frac{3}{4}(c_m + \frac{c_a}{\varepsilon})\right)$ of amplitude ε. It is transported with the velocity $c_m + \frac{c_a}{\varepsilon}$ and the initial condition is given by the solution at time $t = 0$. The computational domain is given by $(0, c_m + \frac{c_a}{\varepsilon})$ with final time $T_f = 1$, which corresponds to one full cycle with periodic boundary conditions. We take $c_m = 1$ and $c_a = 1$.

Since we consider a linear advection problem that respects a maximum principle, the detection criterion Φ in the MOOD Algorithm 1 is given by the L^∞ norm of the solution

$$\Phi(w) = w. \tag{4.4}$$

The hierarchy of the MOOD procedure consists in a high-order scheme given by the third-order IMEX3 or IMEX3(4) schemes from Section 3.3.1 and Section 3.3.2 with centred differences in the implicitly treated space derivative, whereas the parachute scheme is given by the first-order IMEX scheme TVD3 or TVD3(4) with upwind discretization for all derivatives.

We compare our results against the L-stable ARS(2,3,3) scheme, reported in [2] or [34] and recalled in Appendix Appendix C. Note that the ARS(2,3,3) scheme can be written in the form of (3.8) with $\gamma = \frac{3-\sqrt{3}}{6}$, and therefore it falls within the framework of CK schemes that were used in Section 3.1.1. However, this value of γ does not satisfy the requirement of Lemma 1. Consequently, within our framework, we cannot prove the existence of a convex combination, starting with the ARS(2,3,3) scheme, that leads to a first order TVD and L^∞ stable scheme with a material CFL condition (4.2). In the following, we display the numerical results for the IMEX1, ARS(2,3,3), IMEX3, TVD3, MOOD3 and MOOD3(4) schemes.

4.1.1. Maximum principle

We study the behaviour of the above-mentioned schemes with respect to the maximum principle for the discontinuous solution (4.3). First, in Figure 4, the numerical approximations of the discontinuous solution (4.3) for $\varepsilon = 1$ and $\varepsilon = 10^{-3}$ are displayed. For $\varepsilon = 1$, the solutions are computed with a material CFL condition given by $\nu_{\text{mat}} = 0.5$. For $\varepsilon = 10^{-3}$, the results are given in addition using an acoustic CFL condition given by $\nu_{\text{ac}} = 0.5$. These CFL conditions respectively correspond to time steps of $\Delta t_{\text{ac}} \simeq 10^{-5}$

and $\Delta t_{\text{mat}} = 10^{-2}$. In both cases, we take $\Delta x = 0.1$, which corresponds to 20 cells for $\varepsilon = 1$ and 10010 cells for $\varepsilon = 10^{-3}$.

Since the solution contains only fast travelling waves, using a material time step leads to diffused wave fronts. However, for the acoustic time step, they are captured accurately since the CFL condition is limited by these fast waves. We note that the purely third order schemes IMEX3 and ARS$(2,3,3)$ are oscillatory and violate the maximum principle even for $\varepsilon = 1$, even though the ARS$(2,3,3)$ scheme is L-stable. For $\varepsilon = 10^{-3}$, the unlimited third-order schemes are not in-bounds when using the material CFL condition. Their oscillations are not too large thanks to the diffusion from the upwind treatment of the implicit derivative. However, with the acoustic CFL condition, the centred differences on the implicit derivative produce extremely large oscillations which are not displayed in Figure 4. Conversely, the MOOD schemes are in-bounds for both choices of CFL condition and both ε regimes. Note that the IMEX3(4) scheme is more stable and precise than the IMEX3 scheme, which results in a better precision for the MOOD3(4) scheme compared to the MOOD3 scheme. Indeed, the lack of stability of the IMEX3 scheme, especially for small ε, triggers the parachute scheme more often to ensure the maximum principle.

Next, we compute the error on the numerical solution. A standard error computation in the context of finite volume schemes consists in using the L^1 norm, defined by

$$\|w^n\|_1 = \frac{1}{\Delta x} \sum_j |w_j^n|.$$

However, the above norm only measures the average deviation between the exact solution and the numerical approximation. Since we seek to measure of the violation of the maximum principle, we need to take into

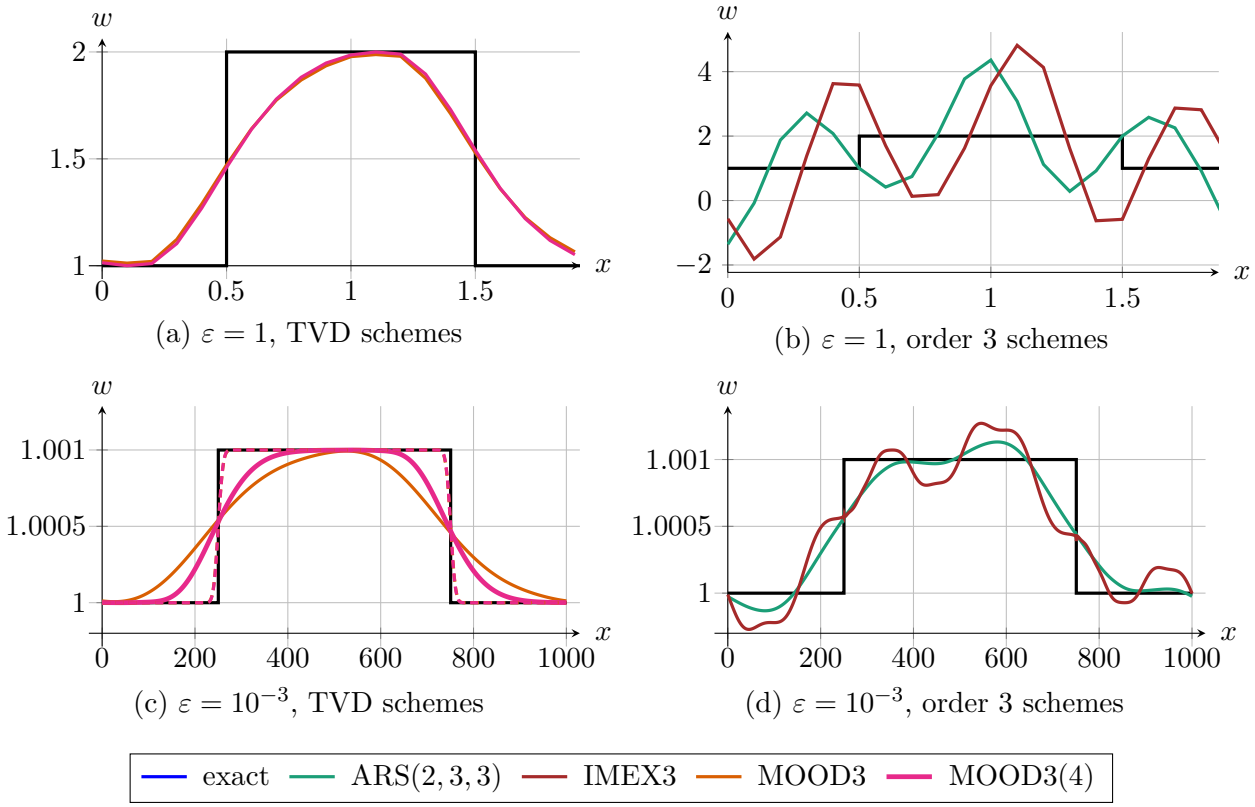


Figure 4: Discontinuous solution (4.3) of the linear advection problem at time $T_f = 1$ with $\Delta x = 0.1$ for $\varepsilon = 1$ (left panel) and $\varepsilon = 10^{-3}$ (middle and right panels). The dashed lines denote the acoustic CFL condition (4.1) with $\nu_{\text{ac}} = 0.5$, and the solid lines use the material CFL condition (4.2) with $\nu_{\text{mat}} = 0.5$.

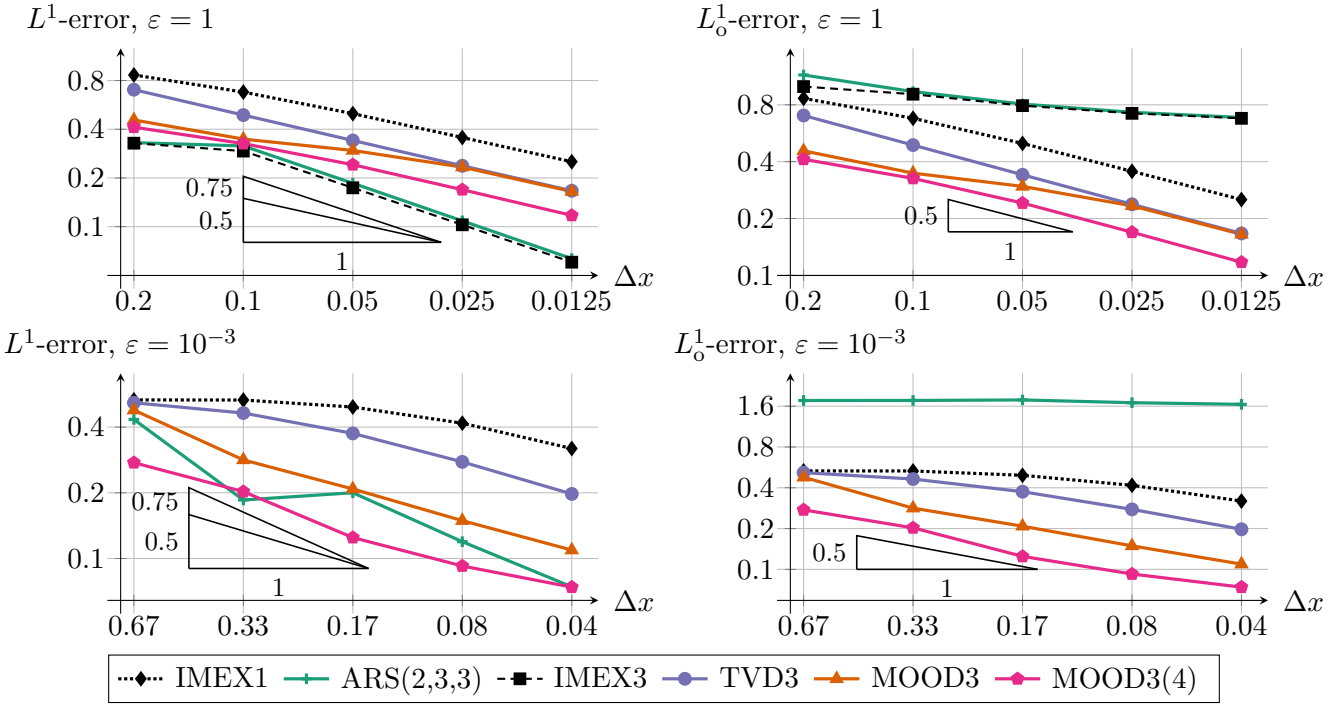


Figure 5: Error lines in L^1 norm (left) and L^1_o quasinorm (right) for the discontinuous solution (4.3) with $\varepsilon = 1$ (top) and $\varepsilon = 10^{-3}$ (bottom).

account the impact of overshoots and undershoots on the error. Therefore, we propose to use the following quasinorm

$$\|w^n\|_{L^1_o} = \frac{1}{\Delta x} \sum_j \left(|w_j^n| + \max_{m \le n} \left[\left(\max_j w_j^m - \min_j w_j^m \right) - \left(\max_j w_j^0 - \min_j w_j^0 \right) \right] \right),$$

which adds the error of the over- and undershoots to the L^1 error.

In Figure 5, we report the error in the L^1 norm and in the L^1_o quasinorm produced by the above-mentioned schemes for $\varepsilon = 1$ and $\varepsilon = 10^{-3}$ with $\nu_{\text{mat}} = 0.5$, corresponding to $\Delta t = 0.01$. First, we observe that the theoretical order of convergence for a higher order scheme is not reached since we obtain an accuracy up to order $\frac{1}{2}$ for the IMEX1, TVD3, MOOD3 and MOOD3(4) schemes, and up to order $\frac{3}{4}$ for the ARS(2,3,3) and IMEX3 schemes. This is due to the fact that we approximate a discontinuous solution where the numerical diffusion of the schemes considerably reduces the order of convergence, see for instance [28]. Second, when taking the over- and undershoots into account, i.e. considering the error in the L^1_o quasinorm, we observe that the IMEX1, TVD3, MOOD3 and MOOD3(4) schemes show comparable experimental order of convergence (EOC) as in the L^1 norm. This was to be expected since they are constructed such that over- or undershoots are avoided. However, for the ARS(2,3,3) scheme, the error in the L^1_o norm is roughly constant even when an increasing number of cells is considered. This means that the improvement of the L^1 error is almost exactly compensated by an increase of the over- and undershoots in absolute magnitude. Therefore, even taking an increasing number of cells is not enough to stabilise the ARS(2,3,3) scheme.

4.1.2. Advantages of the TVD-MOOD approach

We continue our discussion by addressing the flexibility of our schemes with respect to the choice of time stepping compared to L-stable and SSP IMEX schemes from the literature. Further, we show that our first order IMEX-TVD schemes, overall, improve the space-time errors compared to the classical IMEX1 first-order backward forward Euler scheme.

	λ	CPU time (s)	L^1 error
MOOD3(4)	$\lambda = \lambda^{\text{opt}} \simeq 0.548$	0.0101	0.217
	$\lambda = 250\varepsilon = 0.25$	0.0222	0.111
	$\lambda = 50\varepsilon = 0.05$	0.0953	0.0591
	$\lambda = 10\varepsilon = 0.01$	0.659	0.0488
	$\lambda = 2\varepsilon = 0.0002$	1.63	0.0253
	$\lambda = 0.9\varepsilon = 0.0009$	3.64	0.0253
CGGS3	$\lambda = 0.9\varepsilon = 0.0009$	1.25	0.0253
ARS(2,3,3)	$\lambda = 0.9\varepsilon = 0.0009$	1.17	0.0253

Table 2: CPU times and L^1 errors for discontinuous solution (4.3) with $N = 4000$ discretisation points and $\varepsilon = 10^{-3}$, using the MOOD3(4), ARS(2,3,3) and CGGS3.

Flexibility on the time step. In the following, we compare the MOOD3(4) scheme with higher order schemes from the literature. Namely, we consider the aforementioned ARS(2,3,3) scheme from [2], as well as a more recent SSP-IMEX scheme from [8] which we refer to as CGGS3. In contrast to our MOOD3(4) scheme, the ARS(2,3,3) and CGGS3 schemes are restricted by an acoustic CFL condition to be L^∞ stable. This results in a ε-dependent CFL restriction $\lambda \leq 0.9\varepsilon$, which corresponds to $\nu_{ac} = 0.9$ and $\nu_{mat} = 0.009$. Note that our MOOD3(4) scheme allows $\lambda < \lambda^{opt} \simeq 0.547$ corresponding to $\nu_{ac} = 54.7$ and $\nu_{mat} = 0.547$. In Table 2, the CPU time and the L^1 error for the discontinuous solution (4.3) with $N = 4000$ and $\varepsilon = 10^{-3}$ for the MOOD3(4), CGGS3 and ARS(2,3,3) schemes are displayed. Recall that the time step is $\Delta t = \Delta x \frac{\lambda}{1+1/\varepsilon} \simeq 4\lambda$. Applying an acoustic CFL condition gives comparable errors and CPU times for all schemes, leading to a resolution of all waves. However, in the case of our MOOD3(4) scheme, the scheme is still stable for larger values of λ. This can be especially advantageous if the resolution of certain waves can be neglected, which reduces computational time. Examples of this behaviour include the results with a material time step given in Figure 4, or the approximation of material waves in the context of the isentropic Euler equations in Figure 7. Also, note that using high-order schemes such as ARS(2,3,3) or CGGS3 enforces a time step that vanishes if ε tends to zero, leading to long computational times especially when flow phenomena are monitored over long time intervals.

Space-time error of the TVD-MOOD approach. In this paragraph, we study the impact of using the TVD3(4) scheme as a parachute scheme instead of the usual IMEX1 backward forward Euler scheme. For fairness, we compare the TVD3(4) scheme to the four-stage IMEX1(4) version of the IMEX1 scheme. In addition, since we wish to compare the time discretisations, we only consider first-order upwind space discretisations in this paragraph. Note that both the TVD3(4) and IMEX1(4) schemes are first-order accurate. An important point to quantify is the reduction of diffusion reflected in the error made by using our TVD schemes as parachute schemes in the MOOD Algorithm 1. Therefore, we introduce space-time errors, which measure the average and maximum errors between the numerical and exact solution caused by diffusion over the whole computational time. Note that, in contrast, L^1 errors are usually computed at the final time only. The space-time errors are defined as

$$e_{\text{ST}}^{\text{mean}} = \frac{1}{n_t} \sum_{n=1}^{n_t} \left[\left(\max_j (w_{\text{ex}})_j^n - \min_j (w_{\text{ex}})_j^n \right) - \left(\max_j w_j^n - \min_j w_j^n \right) \right],$$

$$e_{\text{ST}}^{\text{max}} = \max_{1 \leq n \leq n_t} \left[\left(\max_j (w_{\text{ex}})_j^n - \min_j (w_{\text{ex}})_j^n \right) - \left(\max_j w_j^n - \min_j w_j^n \right) \right],$$

where n_t is the number of time iterations. In Table 3, we report the values of $e_{\text{ST}}^{\text{mean}}$ and $e_{\text{ST}}^{\text{max}}$ for the discontinuous solution (4.3) with respect to $\varepsilon \in \{1, 10^{-1}, 10^{-2}, 10^{-3}\}$. We take $\Delta x = 0.1$, leading to $N \simeq 20/\varepsilon$ cells since the size of the computational domain depends on ε. We set $\nu_{mat} = 0.5$, which

ε	mean spacetime error			maximum spacetime error		
	🪂: IMEX1(4)	🪂: TVD3(4)	ratio	🪂: IMEX1(4)	🪂: TVD3(4)	ratio
1.0	2.29×10^{-1}	2.29×10^{-1}	1.00	5.20×10^{-1}	5.20×10^{-1}	1.00
10^{-1}	3.41×10^{-3}	3.91×10^{-3}	0.87	1.80×10^{-4}	1.99×10^{-5}	0.90
10^{-2}	3.03×10^{-7}	1.46×10^{-8}	20.8	1.41×10^{-6}	1.30×10^{-7}	10.8
10^{-3}	2.93×10^{-6}	3.21×10^{-8}	91.3	2.37×10^{-5}	2.46×10^{-7}	96.3

Table 3: Space-time error with respect to ε using the TVD3(4) and IMEX1(4) scheme as a parachute for the discontinuous solution (4.3).

corresponds to $\nu_{ac} = \nu_{mat}(c_m + c_a/\varepsilon)/c_m = 0.5 + 0.5/\varepsilon$ and $\Delta t = 0.01$. We note that, for the mean and maximum space-time errors, using the TVD3(4) scheme as a parachute scheme lowers the error by a factor of up to 100 for small values of ε, compared to the IMEX1(4) scheme. This is due to the fact that the TVD3(4) scheme is much less diffusive than the IMEX1(4) scheme for small values of ε, and thus the approximate solution stays closer to the exact solution when the parachute scheme is triggered, rather than being diffused away.

4.2. Isentropic Euler equations

In this section, we apply the TVD-MOOD strategy given by Algorithm 1 to the isentropic Euler equations. They are governed in a non-dimensional formulation by

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \\ \partial_t (\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) + \dfrac{1}{M^2} \nabla p(\rho) = 0, \end{cases} \tag{4.5}$$

where $\rho(x,t) > 0$ denotes the density and $\boldsymbol{u}(x,t)$ the velocity field. Assuming an ideal gas, the pressure law is given by $p(\rho) = \rho^\gamma$, where $\gamma \geq 1$ is the ratio of specific heats. Due to the non-dimensional formulation, the pressure gradient is scaled by the Mach number squared, which is given by the ratio between fluid velocity $\|\boldsymbol{u}\|$ and sound speed c. This leads to Mach number-dependent acoustic wave speeds $\lambda^{\pm} = u \pm \frac{c}{M}$ with $c^2 = \partial_\rho p$. For small Mach numbers, they tend to infinity, and are therefore integrated implicitly. This results in the following splitting of the flux

$$\partial_t w + \nabla \cdot f_e(w) + \nabla \cdot f_i(w) = 0, \tag{4.6}$$

with the state vector $w = (\rho, \rho\boldsymbol{u})^T$, the explicitly treated flux $f_e(w) = (0, \rho \boldsymbol{u} \otimes u)^T$ and the implicitly treated flux $f_i(w) = (\rho\boldsymbol{u}, \frac{1}{M^2} pI)^T$. The reader is referred, for instance, to [26, 10] for a description of the low Mach number limit, and to [9, 4, 43] for more information on the so-called asymptotic-preserving property, which ensures the consistency of the numerical scheme with the incompressible Euler equations in the singular Mach number limit.

Regarding the space discretisation, we consider two schemes given by a RS-IMEX scheme in the spirit of [44, 29] and the TVD IMEX scheme from [12]. Both of them have a Mach number independent CFL condition. The RS-IMEX scheme is obtained by linearising the pressure against a constant reference density, as it appears in the low Mach limit, and the implicitly treated flux is discretised with centred differences. This prevents the scheme from being TVD, but is consistent with the low Mach number limit. In the scheme from [12], all fluxes are discretized in an upwind fashion, which leads to a TVD scheme, but with a numerical viscosity depending on the Mach number, leading to possibly non-feasible numerical solutions in the low Mach number limit. We thus propose the following MOOD hierarchy. First, the high order scheme is the third order IMEX3(4) scheme with the RS-IMEX discretisation. Then, we introduce an intermediate stage with a TVD3(4) integration and RS-IMEX discretization. It has a Mach number-independent diffusion but is overall not TVD due to the centred differences. Finally, the parachute scheme is the TVD3(4) scheme

with the upwind discretisation from [12], which is TVD but very diffusive for low Mach numbers, and should only be triggered rarely.

For the definition of the MOOD criterion (4.4), we follow [11] and use the Riemann invariants to detect oscillations. This choice is motivated by the fact that, according to [40], at least one of the Riemann invariants satisfies a maximum principle in a 1D Riemann problem. It can be extended to the 2D Cartesian framework by using the normal velocity to define the Riemann invariants at each edge of the mesh. The criterion is then given by

$$\Phi_{\pm}(W) = \boldsymbol{u} \cdot \boldsymbol{n} \mp \frac{1}{M} \frac{2}{\gamma - 1} \sqrt{\gamma \rho^{\gamma - 1}}, \tag{4.7}$$

where $\boldsymbol{n}$ is the outward-pointing normal vector. The detection criterion in Algorithm 1 is then time-dependent and given by $\Phi = \max(\Phi_-, \Phi_+)$. Recall from Algorithm 1 that $\xi \in [0, 1]$ controls the permitted oscillations in the MOOD solution. Choosing a small ξ allows the presence of small oscillations in the MOOD solution. However, setting $\xi = 0$ would strictly enforce the detection criterion, and the parachute scheme would be triggered too often, leading to a very diffused solution especially in low Mach number regimes. Therefore, we set $\xi = \frac{1}{100}$ moderately small for all the following numerical experiments. Due to the splitting, the stability of the numerical scheme solely depends on the fluid velocity. The CFL condition in normal direction is thus given by

$$\Delta t \leq \nu_{\text{mat}} \frac{\Delta x}{2 \max |\boldsymbol{u} \cdot \boldsymbol{n}|}, \tag{4.8}$$

which does not depend on the Mach number M. The acoustic CFL condition (4.1) is restricted by $\max |\lambda^{\pm}|$ and enforces a vanishing time step when M tends to 0. In the following numerical tests, we focus on the two-dimensional setting on Cartesian grids.

4.2.1. Riemann problems

We first apply the scheme on two Riemann problems. The first one only concerns the propagation of acoustic waves, while the second one contains a slow moving material wave. For each problem, we provide a reference solution, computed using the IMEX1 scheme on a fine grid. Since the jump in the initial condition is only in x-direction, these tests mimic one-dimensional Riemann problems. Therefore, we consider a mesh with 100 cells in the x-direction and 3 cells in the y-direction. Further, we set $\gamma = 1.4$ and the computational domain is given by $[0, 2] \times [0, 1]$, equipped with Neumann boundary conditions.

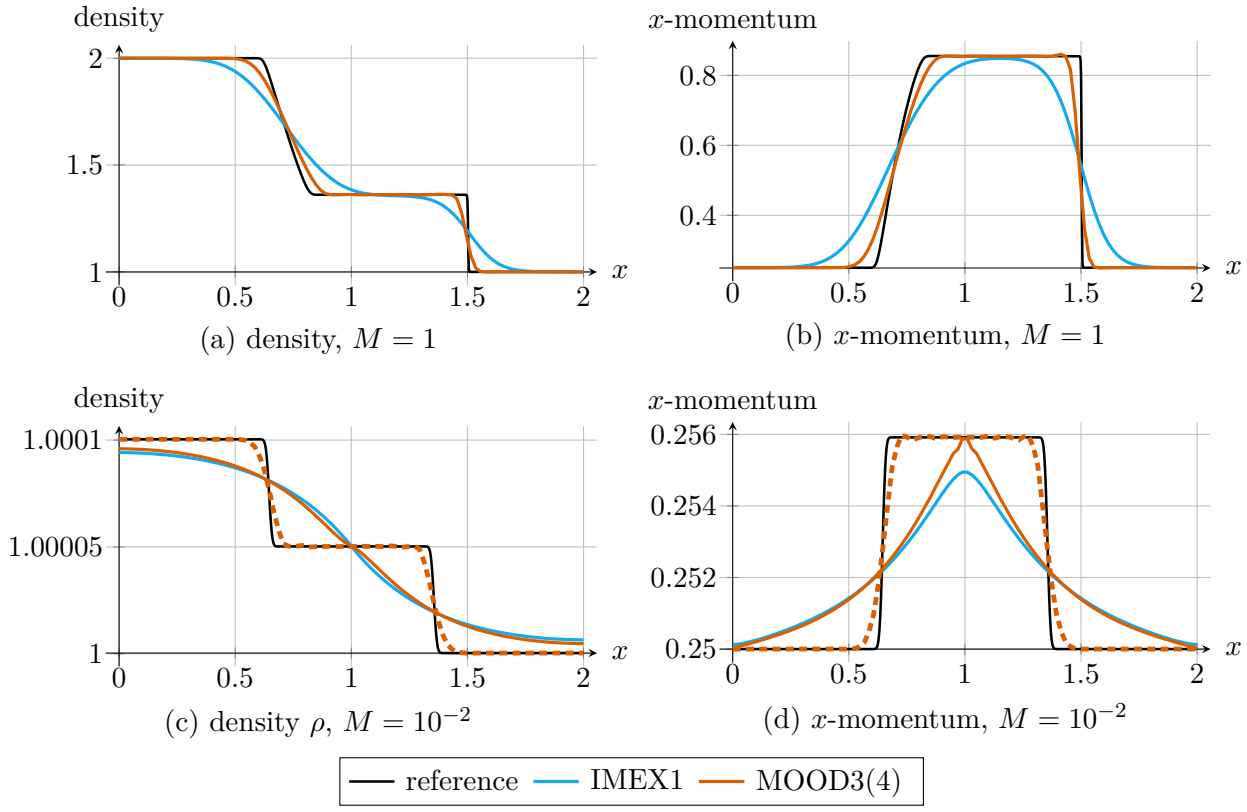
Acoustic waves. The initial data of the Riemann problem is given by

$$\rho(x, y, 0) = \begin{cases} 1 + M^2 & \text{if } x < 1, \\ 1 & \text{otherwise,} \end{cases}, \quad \boldsymbol{u}(x, y, 0) = 0. \tag{4.9}$$

The approximations are depicted at the final time $T_f = 0.3M$ in Figure 6 for $M = 1$ and $M = 10^{-2}$, with $\nu_{\text{ac}} = 0.5$ and $\nu_{\text{mat}} = 0.5$. For $M = 10^{-2}$, these CFL conditions respectively correspond to $\Delta t_{\text{ac}} \simeq 8.43 \times 10^{-5}$ and $\Delta t_{\text{mat}} = 2 \times 10^{-2} > 0.3M$ for $M = 10^{-2}$.

We note that, using the acoustic CFL condition, the shock and rarefaction waves are captured accurately with the correct propagation speed for both Mach numbers. Applying a material CFL condition leads to a stable scheme for $M = 10^{-2}$. As expected for larger time steps, the acoustic waves are diffused but the MOOD3(4) scheme still gives the correct maximal momentum. Utilizing such large time steps allows to get a correct description of the solution shape using only one time step, compared to 36 time steps with an acoustic CFL condition.

Furthermore, for $M = 1$ and $M = 10^{-2}$, the MOOD procedure is activated respectively 51% and 30% of the time iterations using acoustic time stepping. This observation is explained by the fact that the IMEX3(4) scheme is quite oscillatory, which have to be corrected by the MOOD procedure. The share of iterations where the MOOD criterion (4.7) was violated could be lowered by basing the MOOD procedure on a less oscillatory third-order IMEX scheme than the IMEX3(4) scheme, as will be done in Section 4.2.2. Note that, as expected, setting a non zero value of ξ in the MOOD procedure gives rise to marginal oscillations which do not interfere with the quality of the solution.



(a) density, $M = 1$ (b) x-momentum, $M = 1$

(c) density ρ, $M = 10^{-2}$ (d) x-momentum, $M = 10^{-2}$

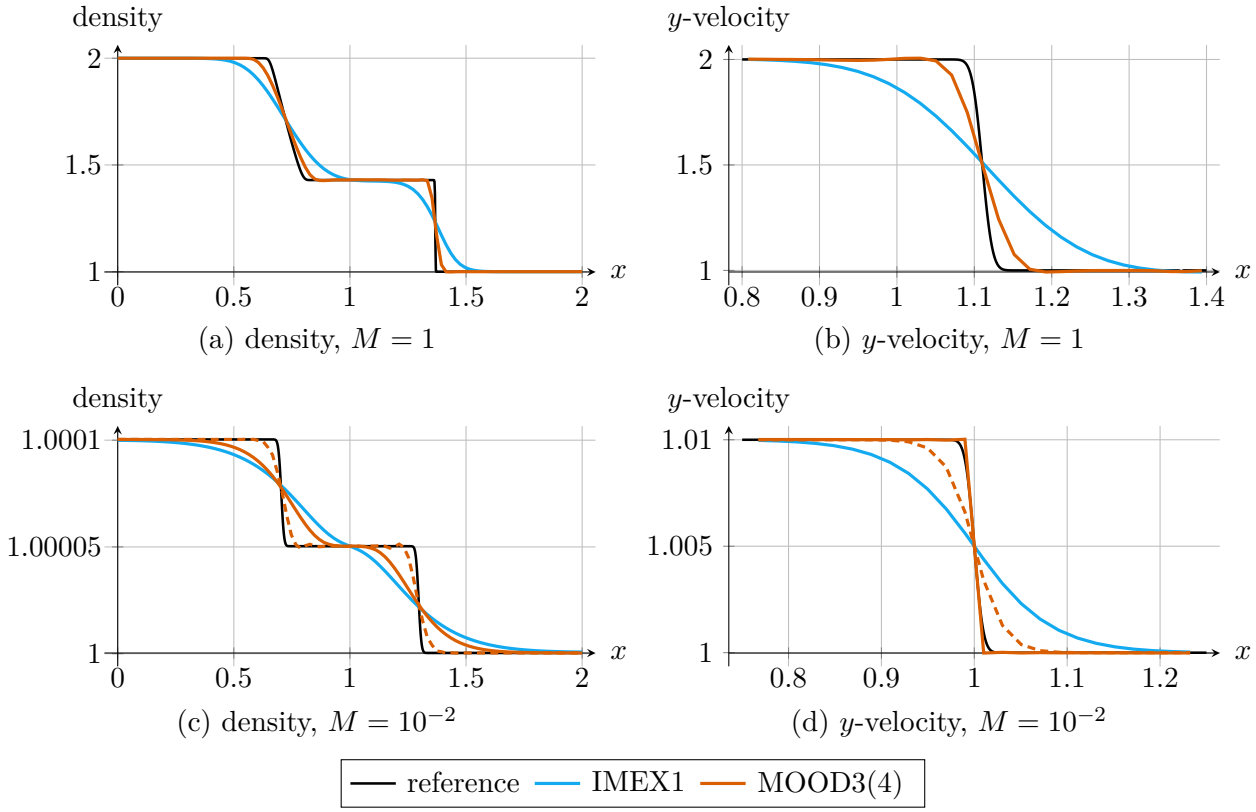
Figure 6: Acoustic waves: Approximation of the solution to the Riemann problem (4.9) at time $T_f = 0.3M$ on a 100×3 mesh, with $M = 1$ (top) and $M = 10^{-2}$ (bottom). For $M = 10^{-2}$, we report the results with material $\nu_{\text{mat}} = 0.5$ (solid lines) and acoustic $\nu_{\text{ac}} = 0.5$ (dashed lines) CFL restrictions.

Material wave. Since the Riemann problem in the previous test consisted only of acoustic waves, we now consider the approximation of a material wave by introducing a shear wave triggered by a non-zero initial velocity in y-direction. The initial condition is given by

$$\rho(x,y,0) = \begin{cases} 1+M^2 & \text{if } x < 1, \\ 1 & \text{otherwise,} \end{cases}, \quad \boldsymbol{u}(x,y,0) = \begin{cases} \begin{pmatrix} 0 \\ 1+M \end{pmatrix} & \text{if } x < 1, \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{otherwise.} \end{cases} \tag{4.10}$$

This experiment verifies that our scheme is able to provide a sharp approximation of the slow-moving shear wave. The numerical results are given in Figure 7 for $M = 1$ and $M = 10^{-2}$, with $\nu_{\text{ac}} = 0.5$ and $\nu_{\text{mat}} = 0.1$, respectively corresponding to $\Delta t_{\text{ac}} \simeq 8.38 \times 10^{-5}$ and $\Delta t_{\text{mat}} = 9.90 \times 10^{-4}$ for $M = 10^{-2}$, at the final time $T_f = 0.25M$.

We observe, as expected, that the approximation of the shear wave is sharp, especially for the material CFL condition. The acoustic waves in the density solution are diffused for the material CFL, which is expected. For the acoustic CFL, all waves are captured with the correct propagation speed. The MOOD procedure is activated respectively 62% and 24% of time iterations for both Mach number regimes using an acoustic CFL condition. Here, the violation of the detection criterion (4.7) leads to the use of the lowest stage of the MOOD procedure, which is based on the TVD time integration and TVD space discretisation.



(a) density, $M = 1$ (b) y-velocity, $M = 1$

(c) density, $M = 10^{-2}$ (d) y-velocity, $M = 10^{-2}$

Figure 7: Shear wave experiment: Approximation of the solution to the Riemann problem (4.10) at time $T_f = 0.25M$ on a 100×3 mesh, with $M = 1$ (top) and $M = 10^{-2}$ (bottom), using the IMEX1 and MOOD3(4) schemes. For $M = 10^{-2}$, we report the results with material $\nu_{\text{mat}} = 0.1$ (solid lines) and acoustic $\nu_{\text{ac}} = 0.5$ (dashed lines) CFL restrictions.

4.2.2. Stationary isentropic vortex

We finally consider a stationary two-dimensional vortex, given by

$$\begin{cases} \rho(x,y) = 1 - \dfrac{M^2}{8} e^{-2a^2 r(x,y)^2}, \\ \boldsymbol{u}(x,y) = a\sqrt{\dfrac{\gamma}{2}} e^{-a^2 r(x,y)^2} \rho(x,y)^{\frac{\gamma}{2}-1} \begin{pmatrix} y \\ -x \end{pmatrix}. \end{cases} \tag{4.11}$$

For the simulation, we set $a = 8$ and consider the computational domain $[0,1]^2$ with periodic boundary conditions up to the final time $T_f = 0.2$. We take $\nu_{\text{mat}} = 0.1$, which corresponds to $\nu_{\text{ac}} \simeq 0.124$ for $M = 1$ and $\nu_{\text{ac}} \simeq 19.2$ for $M = 10^{-2}$. The time step is then given by $\Delta t = 0.01\sqrt{N}/32$.

The L^∞ error lines for $M = 1$ and $M = 10^{-2}$ are reported in Figure 8. We note that the TVD3 scheme is first-order accurate and more precise than the IMEX1 scheme for $M = 1$. For $M = 10^{-2}$, the TVD3(4) and IMEX1 schemes have roughly the same error. This is due to the Mach number-dependent diffusion required required for TVD stability in the numerical scheme, see [12]. This highlights the relevance of our three-stage MOOD algorithm. Further, we note that the MOOD3(4) scheme has the expected EOC of third order. Since the solution is smooth, a reduction of the EOC due to the use of the parachute scheme has not occurred. This confirms the fact that the detection criterion based on the Riemann invariants is able to correctly detect smooth regions, and to appropriately keep the high-order schemes when the solution is oscillation-free. As a consequence, the TVD-MOOD scheme remains third-order accurate for each value of the Mach number M under consideration. Furthermore, we see from the errors that the diffusion is independent of the Mach

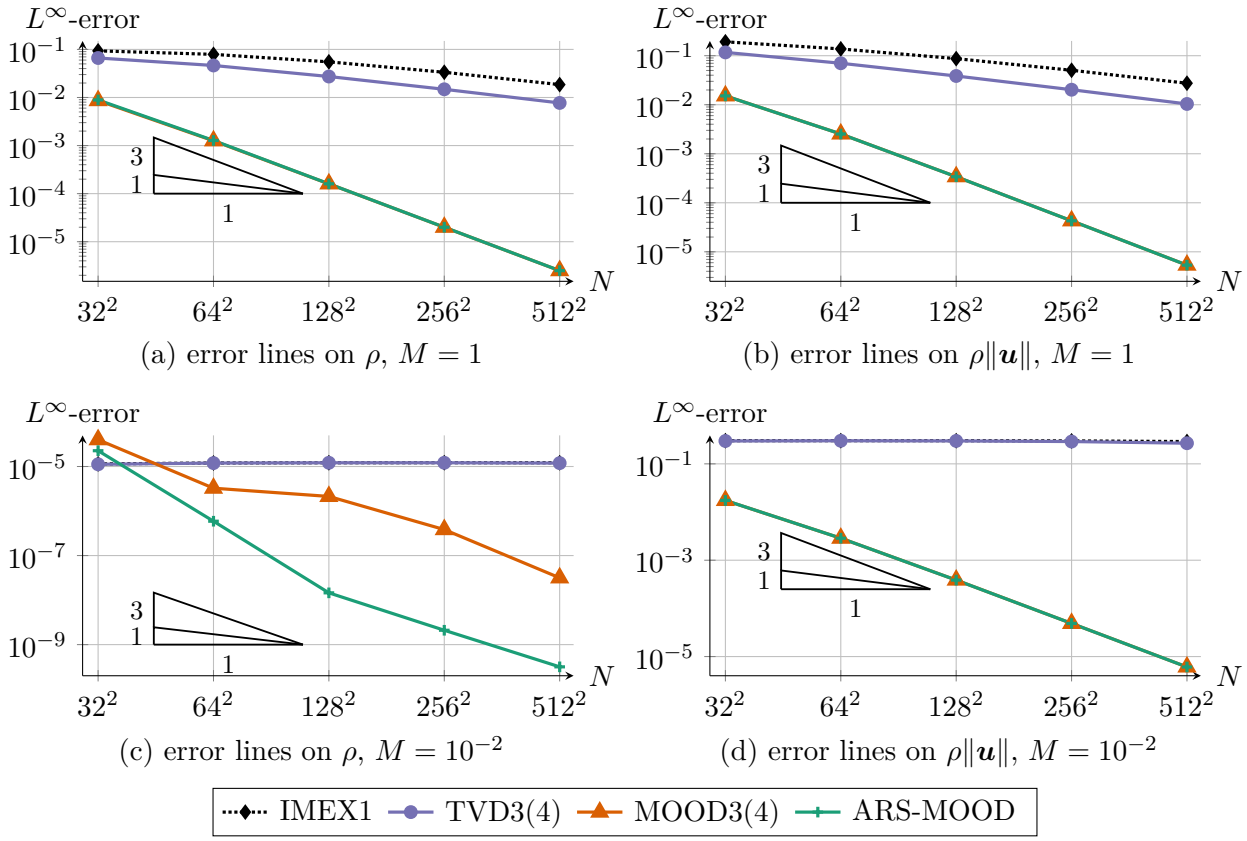


Figure 8: Error lines in L^2 norm for the 2D vortex described in Section 4.2.2, with $M = 1$ (top panels) and $M = 10^{-2}$ (bottom panels). Left panels: errors on the density ρ; right panels: errors on the momentum norm $\rho\|\boldsymbol{u}\|$.

number, and since the initial condition is well-prepared, see [10], it is a strong indication that the scheme is also asymptotic preserving, i.e. is also a consistent discretization of the incompressible Euler equations as M tends to 0.

We also report the error lines of the ARS-MOOD scheme, which consists of performing the MOOD Algorithm 1 with the ARS(2,3,3) as the third order scheme and the TVD3(4) time integrator for the parachute scheme. We note that the ARS-MOOD performs as well as the MOOD3(4) scheme for the errors in the momentum, while it yields significantly better density errors for $M = 10^{-2}$. This can be explained by the increased stability properties of the ARS(2,3,3) leading to a better approximation for small Mach numbers. Although both MOOD3(4) and ARS-MOOD show a third-order EOC, the improved errors of the ARS-MOOD scheme underline the importance of choosing a suitable L-stable or stiffly accurate high order scheme in the MOOD procedure for simulating low Mach flows.

5. Conclusions and future work

We have presented a new approach on constructing first order TVD IMEX-RK schemes, which are especially suited as MOOD parachute schemes to simulate multi-scale equations due to a reduced numerical viscosity compared to a backward forward first order Euler integrator, see Table 3 for a quantification using space-time errors. Their development is motivated by the fact that there is a first order barrier for IMEX-TVD schemes that have a scale independent CFL restriction. The key in constructing those schemes lies in using a convex combination of a first-order TVD IMEX scheme with a high-order IMEX RK scheme. We gave a theoretical justification of our TVD approach by means of studying a one dimensional linear scalar

equation. The obtained TVD scheme is then used as a parachute scheme in a MOOD procedure, which is triggered whenever the TVD property is violated by the high order scheme. The performance of the resulting TVD and MOOD schemes was verified by approximating a discontinuous solution for scalar linear transport, see Figure 4, as well as Riemann problems for the 2D isentropic Euler equations, see Figures 6 and 7 assessing the resolution of acoustic and material waves. Further, it was verified that the MOOD procedure yields high order convergence, correctly detecting smooth solutions, by simulating a vortex in different Mach number regimes, see Figure 8. Our results are a significant improvement to the scheme from [11] for the isentropic Euler equations, especially for small Mach numbers, where our MOOD schemes are able keep a sharp profile on the material wave. Due to the Mach number dependent diffusion of the TVD scheme from [11], the grid would have to be drastically refined to obtain comparable results on the material wave.

The construction of TVD or SSP IMEX schemes with a material CFL restriction for multi-scale equations is still an active field of research, as shown by the recent work of [15]. Therefore, our schemes, which provide a certain flexibility regarding the time step while maintaining the TVD property, are necessary to obtain physically admissible solutions for problems relevant to the community.

Since, in this work, we have neglected a higher order reconstruction of the implicitly treated derivatives due to avoiding the inversion of non-linear systems, in future work we plan to combine our schemes with the linear implicit high order *Quinpi* approach from [36].

Appendix A. On non-CK IMEX schemes

Consider the following Butcher tableaux, defining an IMEX scheme in non-CK, non-ARS form:

$$\text{explicit:}\quad \begin{array}{c|cccc} 0 & 0 & 0 & \cdots & 0 \\ \tilde{c}_2 & \tilde{a}_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \tilde{c}_s & \tilde{a}_{s1} & \cdots & \tilde{a}_{s,s-1} & 0 \\ \hline & \tilde{b}_1 & \cdots & \tilde{b}_{s-1} & \tilde{b}_s \end{array} \qquad \text{implicit:}\quad \begin{array}{c|cccc} c_1 & a_{11} & 0 & \cdots & 0 \\ c_2 & a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array} . \tag{A.1}$$

We derive stability conditions analogous to Theorem 2 for this case where the first column of the implicit tableau is non-zero. After lengthy computations, we get the following result:

Theorem 3. *Let $\tilde{A}, A \in \mathbb{R}^{s\times s}$, $\tilde{b}, b, \tilde{c}, c \in \mathbb{R}^s$ define two Butcher tableaux (A.1) fulfilling (3.4) and the p-th order compatibility conditions. Let $\tilde{b}$ and b coincide with the last rows of $\tilde{A}$ and A respectively. For $k = 1, \dots, s$ and $l = 1, \dots, k-1$, we define*

$$\mathcal{A}_k = \theta_k a_{kk} + (1-\theta_k)c_k, \quad \tilde{\mathcal{A}}_k = (1-\theta_k)\tilde{c}_k, \quad \mathcal{B}_{kl} = \frac{\theta_k a_{kl}}{\mathcal{A}_l}, \quad \tilde{\mathcal{B}}_{kl} = \theta_k \tilde{a}_{kl}.$$

In addition, we recursively define the following expressions:

$$\tilde{\mathcal{C}}_k = \tilde{\mathcal{A}}_k - \sum_{l=2}^{k-1} \mathcal{B}_{kl}\tilde{\mathcal{C}}_l, \qquad \tilde{\mathcal{D}}_{kl} = \tilde{\mathcal{B}}_{kl} - \sum_{r=l+1}^{k-1} \mathcal{B}_{kr}\tilde{\mathcal{D}}_{rl},$$
$$\mathcal{C}_k = 1 - \sum_{l=1}^{k-1} \mathcal{B}_{kl}\mathcal{C}_l, \qquad \mathcal{D}_{kl} = \mathcal{B}_{kl} - \sum_{r=l+1}^{k-1} \mathcal{B}_{kr}\mathcal{D}_{rl}.$$

Then, under the following restrictions for $k = 1, \dots, s$ and $l = 1, \dots, k-1$,

$$\mathcal{A}_k > 0, \qquad 0 \le \lambda\tilde{\mathcal{C}}_k \le \mathcal{C}_k, \qquad 0 \le \lambda\tilde{\mathcal{D}}_{k,l} \le \mathcal{D}_{k,l},$$

the scheme consisting in the convex combination based on the Butcher tableaux (A.1), combined with a TVD limiter, is L^∞ stable and TVD under a CFL condition determined by $\lambda \ge 0$ where λ does not depend on ε.

When performing numerical experiments, we observe that the results of schemes derived under the conditions of Theorem 3 are not as compelling as results of schemes obeying Theorem 2. Therefore, we do not include such schemes in the numerical experiments, but we still state Theorem 3 for the sake of completeness.

Appendix B. TVD3(4)

For the TVD3(4) scheme, the explicit Butcher tableau is given by:

0	0	0	0	0
0.2049503677289891	0.2049503677289891	0	0	0
0.4173127343286904	0.2123925641886599	0.2049201701400305	0	0
0.9048203025659662	−0.4501877125339555	0.3955748607480934	0.9594331543518283	0
	0	0.3354718384287510	0.3487815573407456	0.3157466042305059

,

while the implicit Butcher tableau is given as follows:

0	0	0	0	0
0.2049503677289891	0	0.2049503677289891	0	0
0.4173127343286904	0	0.2040104873103189	0.2133022470183705	0
0.9048203025659662	0	0.3991926529002874	0.4115004113464103	0.0941272383192684
	0	0.3354718384287510	0.3487815573407456	0.3157466042305059

.

Appendix C. ARS(2, 2, 3)

The ARS(2,2,3) scheme from [2] is given by the following tableaux:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \delta & \delta & 0 & 0 \\ 1-\delta & \delta-1 & 2-2\delta & 0 \\ \hline & 0 & 1/2 & 1/2 \end{array}, \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \delta & 0 & \delta & 0 \\ 1-\delta & 0 & 1-2\delta & \delta \\ \hline & 0 & 1/2 & 1/2 \end{array}, \quad \text{with} \delta = \frac{3+\sqrt{3}}{6}.$$

Acknowledgements

This work has been partially funded by a CNRS-INSMI PEPS JCJC project. A.T. has been partially supported by the Gutenberg Research College, JGU Mainz.

References

[1] E. Abbate, A. Iollo, and G. Puppo. An Implicit Scheme for Moving Walls and Multi-Material Interfaces in Weakly Compressible Materials. *Commun. Comput. Phys.*, 27(1):116–144, 2019.

[2] U. M. Ascher, S. J. Ruuth, and R. J. Spiteri. Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. *Appl. Numer. Math.*, 25(2-3):151–167, 1997. Special issue on time integration (Amsterdam, 1996).

[3] G. Bispen, K. R. Arun, M. Lukáčová-Medviďová, and S. Noelle. IMEX large time step finite volume methods for low Froude number shallow water flows. *Commun. Comput. Phys.*, 16(2):307–347, 2014.

[4] S. Boscarino, G. Russo, and L. Scandurra. All Mach Number Second Order Semi-implicit Scheme for the Euler Equations of Gas Dynamics. *J. Sci. Comput.*, 77(2):850–884, 2018.

[5] F. Bouchut, E. Franck, and L. Navoret. A low cost semi-implicit low-Mach relaxation scheme for the full Euler equations. *J. Sci. Comput.*, 83(1):24, 2020.

[6] C. Bresten, S. Gottlieb, Z. Grant, D. Higgs, D. I. Ketcheson, and A. Németh. Explicit strong stability preserving multistep Runge–Kutta methods. *Math. Comput.*, 86(304):747–769, 2017.

[7] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for systems of conservation laws—Multi-dimensional Optimal Order Detection (MOOD). *J. Comput. Phys.*, 230(10):4028–4050, 2011.

[8] S. Conde, S. Gottlieb, Z. J. Grant, and J. N. Shadid. Implicit and Implicit–Explicit Strong Stability Preserving Runge–Kutta Methods with High Linear Order. *J. Sci. Comput.*, 73(2-3):667–690, 2017.

[9] P. Degond and M. Tang. All speed scheme for the low Mach number limit of the isentropic Euler equations. *Commun. Comput. Phys.*, 10(1):1–31, 2011.

[10] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number. *J. Comput. Phys.*, 229(4):978–1016, 2010.

[11] G. Dimarco, R. Loubère, V. Michel-Dansac, and M.-H. Vignal. Second-order implicit-explicit total variation diminishing schemes for the Euler system in the low Mach regime. *J. Comput. Phys.*, 372:178–201, 2018.
[12] G. Dimarco, R. Loubère, and M.-H. Vignal. Study of a New Asymptotic Preserving Scheme for the Euler System in the Low Mach Number Limit. *SIAM J. Sci. Comput*, 39(5):A2099–A2128, 2017.
[13] S. K. Godunov. A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations. *Math. Sbornik*, 47:271–306, 1959.
[14] S. Gottlieb. On high order strong stability preserving Runge-Kutta and multi step time discretizations. *J. Sci. Comput.*, 25(1-2):105–128, 2005.
[15] S. Gottlieb, Z. J. Grant, J. Hu, and R. Shu. High Order Strong Stability Preserving MultiDerivative Implicit and IMEX Runge–Kutta Methods with Asymptotic Preserving Properties. *SIAM J. Numer. Anal.*, 60(1):423–449, 2022.
[16] S. Gottlieb, D. Ketcheson, and C.-W. Shu. *Strong Stability Preserving Runge-Kutta and Multistep Time Discretizations*. WORLD SCIENTIFIC, 2011.
[17] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Math. Comp.*, 67(221):73–85, 1998.
[18] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Rev.*, 43(1):89–112, 2001.
[19] H. Guillard and C. Viozat. On the behaviour of upwind schemes in the low Mach number limit. *Comput. & Fluids*, 28(1):63–86, 1999.
[20] A. Harten. On a Class of High Resolution Total-Variation-Stable Finite-Difference Schemes. *SIAM J. Numer. Anal.*, 21(1):1–23, 1984.
[21] I. Higueras. Strong Stability for Additive Runge–Kutta Methods. *SIAM J. Numer. Anal.*, 44(4):1735–1758, 2006.
[22] I. Higueras, N. Happenhofer, O. Koch, and F. Kupka. Optimized strong stability preserving IMEX Runge–Kutta methods. *J. Comput. Appl. Math.*, 272:116–140, 2014.
[23] I. Higueras, D. I. Ketcheson, and T. A. Kocsis. Optimal Monotonicity-Preserving Perturbations of a Given Runge–Kutta Method. *J. Sci. Comput.*, 76(3):1337–1369, 2018.
[24] X. Y. Hu, N. A. Adams, and C.-W. Shu. Positivity-preserving method for high-order conservative schemes solving compressible Euler equations. *J. Comput. Phys.*, 242:169–180, 2013.
[25] C. A. Kennedy and M. H. Carpenter. Additive Runge–Kutta schemes for convection–diffusion–reaction equations. *Appl. Numer. Math.*, 44(1-2):139–181, 2003.
[26] S. Klainerman and A. Majda. Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids. *Comm. Pure Appl. Math.*, 34(4):481–524, 1981.
[27] R. Klein. Scale-Dependent Models for Atmospheric Flows. *Annu. Rev. Fluid. Mech.*, 42(1):249–274, 2010.
[28] R. J. LeVeque. *Numerical methods for conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 1992.
[29] M. Lukacova-Medvid'ova, G. Puppo, and A. Thomann. An all mach number finite volume method for isentropic two-phase flow. submitted, 2022.
[30] W. H. Matthaeus and M. R. Brown. Nearly incompressible magnetohydrodynamics at low Mach number. *Phys. Fluids*, 31(12):3634, 1988.
[31] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography or Manning friction. *J. Comput. Phys.*, 335:115–154, 2017.
[32] V. Michel-Dansac and A. Thomann. On high-precision L^∞-stable IMEX schemes for scalar hyperbolic multi-scale equations. In *Proceedings of NumHyp 2019*, SEMA SIMAI Springer Series. Springer International Publishing, 2019.
[33] S. Noelle, G. Bispen, K. R. Arun, M. Lukáčová-Medviďová, and C.-D. Munz. A weakly asymptotic preserving low Mach number scheme for the Euler equations of gas dynamics. *SIAM J. Sci. Comput.*, 36(6):B989–B1024, 2014.
[34] L. Pareschi and G. Russo. Implicit-explicit Runge-Kutta schemes for stiff systems of differential equations. In *Recent trends in numerical analysis*, volume 3 of *Adv. Theory Comput. Math.*, pages 269–288. Nova Sci. Publ., Huntington, NY, 2001.
[35] L. Pareschi and G. Russo. Implicit-Explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation. *J. Sci. Comput.*, 25(1-2):129–155, 2005.
[36] G. Puppo, M. Semplice, and G. Visconti. Quinpi: Integrating Conservation Laws with CWENO Implicit Methods. *Commun. Appl. Math. Comput.*, 2022.
[37] P.L. Roe. Generalized formulation of TVD Lax-Wendroff schemes. *ICASE NASA Langley Research Center, Hampton, VA*, ICASE Report No 84-53, 1984.
[38] B. Schmidtmann, B. Seibold, and M. Torrilhon. Relations between WENO3 and third-order limiting in finite volume methods. *J. Sci. Comput.*, 68(2):624–652, 2015.
[39] C.-W. Shu and S. Osher. Efficient implementation of essentially nonoscillatory shock-capturing schemes. *J. Comput. Phys.*, 77(2):439–471, 1988.
[40] J. A. Smoller and J. L. Johnson. Global solutions for an extended class of hyperbolic systems of conservation laws. *Arch. Ration. Mech. Anal.*, 32(3), 1969.
[41] M. N. Spijker. Contractivity in the numerical solution of initial value problems. *Numer. Math.*, 42(3):271–290, 1983.
[42] P. K. Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, 21(5):995–1011, 1984.
[43] A. Thomann, M. Zenk, G. Puppo, and C. Klingenberg. An All Speed Second Order IMEX Relaxation Scheme for the Euler Equations. *Commun. Comput. Phys.*, 28(2):591–620, 2020.
[44] J. Zeifang, J. Schütz, K. Kaiser, A. Beck, M. Lukáčová Medviďová, and S. Noelle. A Novel Full-Euler Low Mach Number IMEX Splitting. *Commun. Comput. Phys.*, 27(1):292–320, 2020.