



**HAL**  
open science

# Second-order uniformly asymptotic-preserving space-time-ImEx schemes for hyperbolic balance laws with stiff relaxation

Louis Reboul, Teddy Pichard, Marc Massot

► **To cite this version:**

Louis Reboul, Teddy Pichard, Marc Massot. Second-order uniformly asymptotic-preserving space-time-ImEx schemes for hyperbolic balance laws with stiff relaxation. 2022. hal-03672571

**HAL Id: hal-03672571**

**<https://hal.science/hal-03672571v1>**

Preprint submitted on 19 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SECOND-ORDER UNIFORMLY ASYMPTOTIC-PRESERVING SPACE-TIME-IMEX SCHEMES FOR HYPERBOLIC BALANCE LAWS WITH STIFF RELAXATION\*

LOUIS REBOUL<sup>†</sup>, TEDDY PICHARD<sup>†</sup>, AND MARC MASSOT<sup>†</sup>

**Abstract.** We consider hyperbolic systems of conservation laws with relaxation source terms leading to a diffusive asymptotic limit under a parabolic scaling. We introduce a new class of second-order in time and space numerical schemes, which are uniformly asymptotic preserving schemes. The proposed Implicit-Explicit (ImEx) approach, does not follow the usual path relying on the method of lines, either with multi-step methods or Runge-Kutta methods, or semi-discretized in time equations, but is inspired from the Lax-Wendroff approach with the proper level of implicit treatment of the source term. As a result, it yields a very compact stencil in space and time and we are able to rigorously show that both the second-order accuracy and the stability conditions are independent of the fast scales in the asymptotic regime, including the study of boundary conditions. We provide an original derivation of  $l^2$  and  $l^\infty$  stability conditions of the scheme that do not deteriorate the second order accuracy without relying on a limiter of any type in the linear case, in particular for shock solutions, and extend such results to the nonlinear case, showing the novelty of the method. The prototype system for the linear case is the hyperbolic heat equation, whereas barotropic Euler equations of gas dynamics with friction are the one for the nonlinear case. The method is also able to yield very accurate steady solutions in the nonlinear case when the relaxation coefficient in the source term depends on space. A thorough numerical assessment of the proposed strategy is provided by investigating smooth solutions, solutions with shocks and solutions leading to a steady state with space dependent relaxation coefficient.

**Key words.** Second order Implicit-Explicit numerical methods, hyperbolic systems of conservation laws with stiff source terms, asymptotic-preserving schemes in diffusive limit,  $l^\infty$ -stable and shock capturing methods

**AMS subject classifications.** 65M06, 65M12, 35L65, 76N15, 76M45, 82C40

**1. Introduction.** We construct and analyze numerical schemes adapted to the hyperbolic-parabolic regimes arising for linear and non-linear hyperbolic problems with a relaxation term. Such systems can result from relaxation approximations of conservation laws [18, 33] as well as fluid limits of kinetic equations [4]. Our incentive originates in the field of plasma physics to perform high-fidelity simulations of sheaths through a new class of numerical methods [3, 35] for fluid models [2]. However, we provide here the key ideas in a simpler context, where numerical analysis is attainable.

We focus on a parabolic scaling corresponding to long-term term behaviors with a stiff source term, and that can be expressed by the introduction of a small parameter  $\varepsilon > 0$ . Consider the system in non-dimensional form:

$$(1.1) \quad \partial_t \mathbf{w} + \frac{1}{\varepsilon} \partial_x \mathbf{f}(\mathbf{w}) = \frac{\sigma}{\varepsilon^2} \mathbf{S}(\mathbf{w})$$

with  $\mathbf{w} \in \Omega \subset \mathbb{R}^n$  a state vector,  $\mathbf{f} : \Omega \mapsto \mathbb{R}^n$  a flux function with a Jacobian matrix  $A(\mathbf{w})$  with a full basis of eigenvalues so that Equation (1.1) is hyperbolic when  $\sigma = 0$ . For the sake of introducing and analyzing a new class of schemes within the scope of this work, we consider only linear source term  $\mathbf{S}(\mathbf{w}) = B\mathbf{w}$  with  $B \in \mathcal{M}_n$ . The asymptotic regime of Equation (1.1) is defined by  $\varepsilon \rightarrow 0$ . The limit depends on

---

\*

**Funding:** The first author is funded by a joint PhD grant from the French Ministry of Defence (AID - Agence Innovation Défense) and the Région Île-de-France (DIM MathInnov).

<sup>†</sup>CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91128 Palaiseau Cedex, France ([louis.reboul@polytechnique.edu](mailto:louis.reboul@polytechnique.edu)).

the specific flux  $\mathbf{f}$  and source term  $\mathbf{S}$ . In this paper, we investigate both a linear case: the classical hyperbolic heat equation (HHE), and a non-linear case: the Euler equations with friction (Euler-friction). These two models are representative of the main difficulties that arise when trying to approximate solutions of equations of the form (1.1). As it is well known (see e.g. [23, 13]), the disparity of scales renders the classical Finite-Volume methods overly diffusive and consequently poorly accurate in such regimes, whereas stability conditions become exceedingly restrictive.

Preserving an accurate description of the asymptotic phenomena with good stability properties while discretizing the original PDE has been the subject of a vast body of literature within the framework of so-called asymptotic-preserving (AP) schemes. Several clever strategies have been designed: combination of source and flux terms [20, 24, 5], control of the numerical dissipation [17], relaxation techniques when the stiffness of the problem is carried by linear operators [25], ImEx strategies in order to benefit from stability conditions close to that of fully implicit schemes but at the cost of explicit computations [7]. Several contributions, as the present one, focus on the parabolic scaling specifically [29, 24, 25, 26].

We focus here on ImEx techniques, that have shown to be suitable for the design of high-order in time and space AP-methods (see [1] and ref. therein) over a wide range of application. However, their construction usually focuses on the time discretization and tend to overlook the challenges linked to the spatial approximations. For instance, stability analysis is routinely performed under the hypothesis that the discrete spatial operator is exact (so-called semi-discretization in time). As a result, such methods may lack of suitable spatial discretizations and sometime exploit unnecessarily complex methods. Similarly, truncation error studies are customarily replaced by simpler procedure, such as verifying only that the limit scheme when  $\varepsilon \rightarrow 0$  produces an approximation of the desired order of the limit system, and potentially missing accuracy loss in intermediate regimes. Current ImEx methods decouple time and space, either using linear-multistep (LM, [1]) or Runge-Kutta (RK, [7]) methods, which require large stencils and, consequently, complex boundary and initial conditions are used. Similar conclusions can be drawn for other methods (e.g. [17]).

The purpose of the present contribution is to provide a class of numerical schemes that are high-order in space and time, uniformly asymptotic preserving in a diffusive limit, and applicable to linear and non-linear systems of conservation laws with stiff source terms. It is designed by coupling intrinsically time and space, as in the Lax-Wendroff approach [31] for systems of conservation laws, but in the presence of source [41]. Such an approach does not rely on a classical ODE approach or semi-discretized equations and it yields an ImEx strategy with a very compact stencil in space and time. It has the ability to deal properly with boundary conditions and offers the possibility to conduct fine stability and truncation error analyses. In this paper, we stick to second order accuracy and only aim at providing stability constraints, which are uniform in the small parameter  $\varepsilon$ . But we do not go fully implicit on the (potentially) non-linear fluxes. Thus, we do not alleviate the natural parabolic stiffness limitation of explicit methods for diffusion equations as in [7, 8, 6]. However, we introduce a novel implicit treatment of the source (Reverse Runge-Kutta methods presented in Appendix C) and provide an original derivation of  $l^2$  and  $l^\infty$  stability conditions of the uniformly AP second order scheme in space and time. Thus, it yields the ability to treat shocks without resorting to limiters in the linear case. For the first and second order schemes applied to the HHE, we conduct a detailed truncation error and stability analysis, clarifying which spatial operator are the viable options compatible with the chosen discrete time structure. We also carefully consider the

question of boundary conditions and provide numerical test-cases that involve non-trivial boundary conditions. Our method is then applied to Euler-friction. To our knowledge, no second order method both in time and space for this set of equations was provided in the literature. We eventually present a numerical assessment of the new scheme with a thorough numerical study of smooth and shock solutions for the HHE, and of smooth, shock solutions as well as a stationary solution with variable in space relaxation coefficient, showing the ability of the method to capture accurately the steady state, even if not formally well-balanced (in the sense of [9]).

The paper is organized as follows: [Section 2](#) presents the models and asymptotic limits we will investigate in the present work. [Section 3](#) introduces relevant numerical notions as well as recalls the limitations of classical finite volume approach. First and second order AP ImEx methods are introduced for the HHE and thoroughly studied in [Sections 4](#) and [5](#). Methods are then extended to the nonlinear case of Euler-friction in [Section 6](#) and verified numerically in [Section 7](#).

**2. Models.** In what follows, we introduce the HHE, a linear model on which the new methods will be derived and their numerical analysis conducted, as well as the Euler-friction equations that will stand as our prototype nonlinear model.

**2.1. Hyperbolic heat equations.** The system of HHE [40, 15, 16, 32] reads:

$$(2.1a) \quad \partial_t E + \frac{1}{\varepsilon} \partial_x F = 0,$$

$$(2.1b) \quad \partial_t F + \frac{1}{\varepsilon} \partial_x E = -\frac{\sigma}{\varepsilon^2} F.$$

We do not detail here the various related physics (propagation of heat at finite speed, low-order moment model for radiative transfer modeling [10], or electrical line transmission so-called the telegrapher's equation).

**2.1.1. Goldstein-Taylor model.** An equivalent form of [Equation \(2.1\)](#) can be found in setting  $u = E + F$ ,  $v = E - F$  to obtain the Goldstein-Taylor model [20]:

$$(2.2a) \quad \partial_t u + \frac{1}{\varepsilon} \partial_x u = -\frac{\sigma}{\varepsilon^2} \frac{u - v}{2},$$

$$(2.2b) \quad \partial_t v - \frac{1}{\varepsilon} \partial_x v = -\frac{\sigma}{\varepsilon^2} \frac{v - u}{2}.$$

This equivalent formulation is particularly convenient to study the stability of numerical methods developed in this paper. The variables  $u$  and  $v$  can also be seen as the Riemann invariants of the convective part of equations (2.1).

**2.1.2. Boundary conditions.** We consider a finite domain of simulation  $I = [x_L, x_R]$ , with  $x_L, x_R \in \mathbb{R}$ ,  $x_L < x_R$ . The question of boundary conditions naturally arises in this setting. We will consider two types of boundary conditions:

- *Periodic* boundary conditions:  $E(x_L) = E(x_R)$ ,  $F(x_L) = F(x_R)$ .
- *Hybrid Dirichlet-Neumann* boundary conditions: imposing a constant Dirichlet boundary conditions for  $E$ , the structure of [Equation \(2.1\)](#) imposes  $\partial_x F = -\varepsilon \partial_t E = 0$ , that is homogeneous Neumann boundary conditions on  $F$ .

To our knowledge this second case is seldom detailed in the literature, although it appears indirectly in some exact test-cases used, e.g. in [17, 10].

**2.1.3. Energy dissipation.** If we consider a smooth solution of [Equation \(2.1\)](#), we can obtain by multiplying [Equation \(2.1a\)](#) by  $E$  and [Equation \(2.1b\)](#) by  $F$ :

$$(2.3) \quad \partial_t (E^2 + F^2) + \frac{2}{\varepsilon} \partial_x (EF) = -\frac{2\sigma}{\varepsilon^2} F^2.$$

Assuming periodic or hybrid homogeneous Dirichlet-Neumann boundary conditions and integrating relation (2.3) over  $\Omega$  we obtain that:

$$(2.4) \quad \partial_t \int_{\Omega} (E^2 + F^2)(t, x) dx \leq 0.$$

We will expect our numerical schemes to obey a discrete equivalent of (2.4), namely it should feature some form of  $l^2$ -stability.

**2.1.4. Asymptotic regime.** A general method to study the asymptotic limit of a model of the form (1.1) is to use a Chapman-Enskog like approach, see e.g. [17]. It consists in considering the following formal expansion of the state vector  $\mathbf{w}$ :

$$(2.5) \quad \mathbf{w} = \mathbf{w}_0 + \varepsilon \mathbf{w}_1 + \mathcal{O}(\varepsilon^2).$$

Injecting this expansion into Equation (1.1) yields a hierarchy of equations that customarily leads to the limit equation at order zero, that is a close-form equation on  $\mathbf{w}_0$ . Using expansion (2.5) into (2.1b) yields  $F_0 = 0$  (at order  $\mathcal{O}(\varepsilon^{-2})$ ) and  $\sigma F_1 = \partial_x E_0$  (order  $\mathcal{O}(\varepsilon^{-1})$ ), which ultimately leads to the diffusive limit:

$$(2.6) \quad \partial_t E_0 - \partial_x \left( \frac{1}{\sigma} \partial_x E_0 \right) = 0,$$

when injected into Equation (2.1a) at order  $\mathcal{O}(\varepsilon^0)$ . This relatively simple linear system already remarkably embodies the notion of asymptotic regime, in that it behaves as an hyperbolic system when  $\varepsilon \sim 1$  but degenerates to a parabolic diffusive limit in the regime  $\varepsilon \rightarrow 0$ . As such it encompasses a significant part of the challenges that arise when designing numerical methods for systems of the form (1.1) while being linear and therefore making the study of accuracy and stability properties for newly elaborated schemes relatively amenable.

**2.2. Euler-friction.** The second model considered in this paper is that of barotropic Euler equations with high friction:

$$(2.7a) \quad \partial_t \rho + \frac{1}{\varepsilon} \partial_x (\rho u) = 0,$$

$$(2.7b) \quad \partial_t (\rho u) + \frac{1}{\varepsilon} \partial_x (\rho u^2 + p(\rho)) = -\frac{\sigma}{\varepsilon^2} \rho u,$$

where  $\rho$  is the density of the fluid and  $u$  its macroscopic velocity and where we assume that pressure verifies  $p'(\rho) > 0$  [5]. As compared to HHE, Equation (2.7) introduces all the difficulties linked with nonlinear hyperbolic systems, such as the onset of shocks in finite time from regular initial data or the necessity to preserve the invariance of the convex space of admissible states  $\{(\rho, \rho u) \in \mathbb{R}^2, \rho \geq 0\}$ .

**2.2.1. Linearized equations.** It is possible to linearize (2.7) around a solution  $\rho_0, (\rho u)_0$ , considering the perturbed solution  $\rho = \rho_0 + \tilde{\rho}$  and  $(\rho u)_0 + \tilde{\rho} u$  with  $\tilde{\rho} \ll \rho_0$  and  $\tilde{\rho} u \ll (\rho u)_0$ , leading to the equations:

$$(2.8a) \quad \partial_t \tilde{\rho} + \frac{1}{\varepsilon} \partial_x \tilde{\rho} u = 0,$$

$$(2.8b) \quad \partial_t \tilde{\rho} u + \frac{1}{\varepsilon} \partial_x (2u_0 \tilde{\rho} u + (c_0^2 - u_0^2) \tilde{\rho}) = -\frac{\sigma}{\varepsilon^2} \tilde{\rho} u,$$

where  $u_0 = (\rho u)_0 / \rho_0$  and  $c_0 = p'(\rho_0)$ . One can notice that when  $u_0 = 0$  and  $c_0 = 1$  equations (2.8) reduce to equations (2.1), by setting  $E = \tilde{\rho}$  and  $F = \tilde{\rho} u$ .

**2.2.2. Asymptotic regime.** Using an analogous approach as the one used for the HHE (2.1), one can derive the diffusive limit when  $\varepsilon \rightarrow 0$ , see e.g. [5]:

$$(2.9) \quad \partial_t \rho_0 - \partial_x \left( \frac{1}{\sigma} \partial_x (p(\rho_0)) \right) = 0,$$

which in the isothermal case and with the law of perfect gases, that is  $p(\rho) = c^2 \rho$ ,  $c > 0$ , is nearly identical to the diffusive limit (2.6).

**3. Numerical basic notions and notations.** The new methods are derived and their properties studied on the HHE. Consequently in this section we introduce, in a linear framework, the notations used in the rest of this paper. We recall the notions of stability, truncation and consistency errors for linear numerical schemes.

**3.1. Numerical approximation.** Let  $w$  be a solution of Equation (1.1) over  $I \times \mathbb{R}_+$ . Finite-volume methods aim at approximating the average of  $w$  over the cells  $C_j = [x_{j+1/2}, x_{j-1/2}]$  at time  $t^n = t_0 + n\Delta t$ :

$$(3.1) \quad \bar{w}_j^n = \frac{1}{|C_j|} \int_{C_j} w(t^n, x) dx,$$

for  $1 \leq j \leq N$  and  $n \geq 0$ , where  $x_{j+1/2} = j\Delta x + x_L$  with a mesh size  $\Delta x = (x_R - x_L)/(N+1)$  and a time step  $\Delta t > 0$  both assumed constant to alleviate notations. The approximation of  $\bar{w}_j^n$  is usually denoted  $w_j^n$ , and the accuracy of a numerical method is evaluated via its *global error*:  $e_j^n = \bar{w}_j^n - w_j^n$ . A method is said to be of order  $p$  in time and  $q$  in space if  $\|e_j^n\| = \mathcal{O}(\Delta t^p + \Delta x^q)$  where  $\|\cdot\|$  is a suitable norm depending on the expected regularity of the solutions. Here we will consider the norms  $\|\cdot\|_\infty$  and  $\|\cdot\|_2$ .

**3.2. Upwind scheme.** Our prototype linear finite-volume scheme is derived using upwind fluxes and centered source terms for (2.2) [10]. For HHE (2.1), it reads:

$$(3.2a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} - \frac{\Delta x}{2\varepsilon} \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2} = 0,$$

$$(3.2b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{E_{j+1}^n - E_{j-1}^n}{2\Delta x} - \frac{\Delta x}{2\varepsilon} \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\Delta x^2} = -\frac{\sigma}{2\varepsilon^2} F_j^n.$$

On this simple linear model all classical finite volume approximated fluxes (upwind, Roe, Rusanov, HLL) yield the exact same scheme, hence our choice of the upwind scheme as an example and reference to be later compared with AP-schemes.

**3.3. Boundary conditions.** The computation of  $E_1^{n+1}$ ,  $E_N^{n+1}$ ,  $F_1^{n+1}$  and  $F_N^{n+1}$  in the scheme (3.2) involves the values  $E_0^{n+1}$ ,  $E_{N+1}^{n+1}$ ,  $F_0^{n+1}$  and  $F_{N+1}^{n+1}$  in ghost-cells. To impose periodic boundary conditions the natural choice is:

$$(3.3) \quad E_0^n = E_N^n, \quad F_0^n = F_N^n, \quad E_{N+1}^n = E_1^n, \quad F_{N+1}^n = F_1^n.$$

A second order discrete form of the hybrid Dirichlet-Neumann boundary conditions presented in Subsection 2.1.2 reads:

$$(3.4) \quad \frac{E_0^n + E_1^n}{2} = E_L, \quad \frac{F_1^n - F_0^n}{\Delta x} = 0, \quad \frac{E_N^n + E_{N+1}^n}{2} = E_R, \quad \frac{F_{N+1}^n - F_N^n}{\Delta x} = 0,$$

where  $E_L$  and  $E_R$  are the values imposed on  $E$  at the left and right boundaries.

Both these sets of boundary conditions are compatible with any scheme for the hyperbolic heat equations (2.1) that are presented in this paper.

**3.4. Convergence.** In the linear framework, convergence is obtained through Lax theorem (see e.g. [30]) by showing that the scheme is consistent and stable.

**3.4.1. Stability.** This discretization of (2.1) yields the  $l^2$ -*diminishing* property:

$$(3.5) \quad \sum_j \left( (E_j^{n+1})^2 + (F_j^{n+1})^2 \right) \Delta x \leq \sum_j \left( (E_j^n)^2 + (F_j^n)^2 \right) \Delta x.$$

This automatically grants the  $l^2$ -*stability*, that is:

$$\exists K > 0, \forall n \geq 0, \quad \sum_j \left( (E_j^n)^2 + (F_j^n)^2 \right) \Delta x \leq K \sum_j \left( (E_j^0)^2 + (F_j^0)^2 \right) \Delta x.$$

**THEOREM 3.1.** *The upwind scheme (3.2) is  $l^2$ -diminishing if the condition:*

$$(3.6) \quad \frac{\Delta t}{\varepsilon \Delta x} + \frac{\sigma \Delta t}{\varepsilon^2} \leq 1$$

*is satisfied or equivalently  $\Delta t \leq \Delta t_{max}$  with  $\Delta t_{max} = (\varepsilon^2 \Delta x) / (\varepsilon + \sigma \Delta x)$ .*

See [10] for a proof of this result. This very strict stability constraint can be alleviated a little by using an implicit source term but remains of the order  $\Delta t_{max} = \mathcal{O}(\varepsilon \Delta x)$ .

**3.4.2. Consistency.** Consistency is studied via truncation error analysis. The *truncation error* is defined by  $\varepsilon_j^n = \bar{w}_j^{n+1} - w_j^{n+1}$ , where  $w_j^{n+1}$  is the approximation yielded by the scheme at time  $t^{n+1}$  under the assumption that  $w_j^n = \bar{w}_j^n$  is exact at previous time  $t^n$ . For a stable scheme, the truncation error is linked to the global error via a relation of the form  $\|e_j^n\| \leq C \sum_{k=0}^n \|\varepsilon_j^k\|$ . Consequently, we need  $\varepsilon_j^n = \Delta t \mathcal{O}(\Delta t^p + \Delta x^q)$  for the scheme to be order  $p$  and  $q$  respectively in time and space.

An equivalent notion that allows to alleviate computations is the *consistency error*  $c_j^n = \varepsilon_j^n / \Delta t$ . Formally, the consistency error can be seen as the error obtained by injecting the exact solution of problem (2.1) into the numerical scheme (3.2).

Lastly, as smooth solutions are considered in consistency studies, we use the equivalent formalism of finite difference for the sake of legibility. Basically, we replace  $\bar{w}_j^n$  by  $w(t^n, x_j)$  with  $x_j = (j + 1/2) \Delta x$  in error terms.

**THEOREM 3.2.** *The consistency error for the upwind scheme (3.2) satisfies:*

$$(3.7) \quad c_j^n(E) = \mathcal{O} \left( \Delta t + \frac{\Delta x}{\varepsilon} \right), \quad c_j^n(F) = \mathcal{O} \left( \Delta t + \Delta x + \frac{\Delta x^2}{\varepsilon} \right).$$

Hence, the global error of the upwind scheme scales as  $e_j^n = \mathcal{O}(\Delta t + \Delta x(1 + 1/\varepsilon))$ . The proof is similar to the one proposed in [10]. We conclude with a few remarks:

- **Theorems 3.1** and **3.2** assess the claim that classical finite volume methods are inaccurate and come at high computational cost in the diffusive regime.
- From now on, instead of considering regimes defined by  $\varepsilon \rightarrow 0$  or  $\varepsilon \sim 1$ , as in the continuous case, we rather consider the regimes  $\varepsilon \ll \Delta x$ ,  $\varepsilon \sim \Delta x$  and  $\Delta x \ll \varepsilon$ , more suitable to the discrete formalism.

**4. First-order AP-schemes.** We first present our formalism, that paves the way for the second order and nonlinear extensions of the AP-methods, on a first order AP-scheme applied to HHE (2.1). At order one, various time integration methods (Runge-Kutta, multi-step, Taylor series LW) are nearly interchangeable and consequently this formalism is similar to the one used in [1, 7].

**4.1. Principle.** Let us write the general structure of the scheme we consider:

$$(4.1a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_x F]_j^{n+1} = 0,$$

$$(4.1b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_x E]_j^{n+1} = -\frac{\sigma}{\varepsilon^2} F_j^{n+1},$$

where  $[\partial_x E]_j^{n+1}$  and  $[\partial_x F]_j^{n+1}$  must be consistent approximations of  $\partial_x E$  and  $\partial_x F$  at time  $t^{n+1}$  and position  $x_j$ . This structure encompasses an implicit version of Scheme (3.2) where fluxes and source are approximated at time  $t^{n+1}$  instead of  $t^n$ . The reason for approximating  $\partial_x F$  at time  $t^{n+1}$  is that it features some fast scales and needs to be implicit in order to obtain a CFL stability condition less restrictive than that of the explicit upwind scheme.

The main difference with other ImEx schemes found in the literature [1, 7] is that we consider also  $[\partial_x E]_j^{n+1}$  instead of  $[\partial_x E]_j^n$  in Equation (4.1b). We did so for two reasons: to make the scheme more symmetric, which will make stability analysis significantly simpler, and to anticipate the fact that any part of the flux can contain fast scales in the general case (1.1), and more particularly for (2.7).

A fully implicit scheme would come at great computational cost as one would have to solve linear systems. The flux terms in (4.1) are defined through a formal discrete equivalent of the partial derivative with respect to space of Equation (2.1):

$$(4.2a) \quad \frac{[\partial_x E]_j^{n+1} - [\partial_x E]_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_{xx} F]_j^n = 0,$$

$$(4.2b) \quad \frac{[\partial_x F]_j^{n+1} - [\partial_x F]_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_{xx} E]_j^n = -\frac{\sigma}{\varepsilon^2} [\partial_x F]_j^{n+1}.$$

with explicit diffusion terms and an implicit source term. The reason for this choice is in the strength of this ImEx approach: it comes at the cost of a scalar linear implicit scheme, that is almost at explicit cost, but with significantly improved stability properties.

Injecting Equation (4.2) into Equation (4.1), we obtain:

$$(4.3a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M}{\varepsilon} [\partial_x F]_j^n - \frac{M\Delta t}{\varepsilon^2} [\partial_{xx} E]_j^n = 0,$$

$$(4.3b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{M}{\varepsilon} [\partial_x E]_j^n - \frac{M\Delta t}{\varepsilon^2} [\partial_{xx} F]_j^n = -\frac{\sigma M}{\varepsilon^2} F_j^n,$$

where  $M = 1/(1 + (\sigma\Delta t)/\varepsilon^2)$ . This factor  $M$  embodies the effect of our implicit step (4.2) at the cost of an explicit step. When  $\Delta t \ll \varepsilon$ , it behaves as a perturbation of one of order one in  $\Delta t$ :  $M = 1 + \mathcal{O}(\Delta t)$ . However, in regimes where  $\varepsilon \ll \Delta t$ , it follows a scaling  $M = \mathcal{O}(\varepsilon^2)$  so the  $M$  factor will help control terms that scale as  $\varepsilon^{-1}$  or  $\varepsilon^{-2}$ .

We will systematically discretize the second order derivative in space terms using the classical centered approximations:

$$[\partial_{xx} E]_j^n = \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2}, \quad [\partial_{xx} F]_j^n = \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\Delta x^2}.$$

The choice of discretization for  $[\partial_x E]_j^n$  and  $[\partial_x F]_j^n$  is critical to recover uniform accuracy with respect to  $\varepsilon$ . We can use for instance the upwind or centered discretization. Ultimately our choice lie in a trade-off between accuracy and stability. In what



follows, we discuss such options, showing that surprisingly the centered approximation is a viable option while using an upwind approximation of the fluxes, as in (3.2), is not accurate in intermediate regimes where  $\Delta x \sim \varepsilon$  (this case is detailed in Appendix B).

**4.2. Centered discretization of the fluxes.** If we use center discretization:

$$[\partial_x E]_j^n = \frac{E_{j+1}^n - E_{j-1}^n}{2\Delta x}, \quad [\partial_x F]_j^n = \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x},$$

we obtain the following first order scheme, afterward called ImEx1-ctr scheme:

$$(4.4a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M}{\varepsilon} \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} - \frac{M\Delta t}{\varepsilon^2} \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2} = 0,$$

$$(4.4b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{M}{\varepsilon} \frac{E_{j+1}^n - E_{j-1}^n}{2\Delta x} - \frac{M\Delta t}{\varepsilon^2} \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\Delta x^2} = -M \frac{\sigma}{\varepsilon^2} F_j^n.$$

#### 4.2.1. $l^\infty$ -stability.

**THEOREM 4.1.** *The ImEx1-ctr scheme (4.4) together with periodic (3.3) or hybrid (3.4) boundary conditions is  $l^\infty$ -diminishing for the variables  $u, v$ , that is:*

$$(4.5) \quad \max_{1 \leq j \leq N} (|u_j^{n+1}|, |v_j^{n+1}|) \leq \max_{1 \leq j \leq N} (|u_j^n|, |v_j^n|),$$

under the condition:

$$(4.6) \quad \Delta t_{min} := \frac{\varepsilon \Delta x}{2} \leq \Delta t \leq \Delta t_{max} := \frac{\sigma \Delta x^2}{4} \frac{1 + \sqrt{1 + 2 \left( \frac{4\varepsilon}{\sigma \Delta x} \right)^2}}{2}.$$

*Proof.* To conduct the stability analysis, we switch to variables  $u = E + F$  and  $v = E - F$ , for which the scheme (4.4) can be rewritten:

$$(4.7a) \quad u_j^{n+1} = \lambda_1 u_j^n + \lambda_2 u_{j+1}^n + \lambda_3 u_{j-1}^n + \lambda_4 v_j^n, \quad \lambda_1 = 1 - \frac{2M\Delta t^2}{\varepsilon^2 \Delta x^2} - \frac{M\sigma \Delta t}{2\varepsilon^2},$$

$$(4.7b) \quad v_j^{n+1} = \lambda_1 v_j^n + \lambda_2 v_{j-1}^n + \lambda_3 v_{j+1}^n + \lambda_4 u_j^n, \quad \lambda_2 = \frac{M\Delta t^2}{\varepsilon^2 \Delta x^2} - \frac{M\Delta t}{2\varepsilon \Delta x},$$

$$\lambda_3 = \frac{M\Delta t^2}{\varepsilon^2 \Delta x^2} + \frac{M\Delta t}{2\varepsilon \Delta x}, \quad \lambda_4 = \frac{M\sigma \Delta t}{2\varepsilon^2}.$$

One notices that  $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$ . Therefore we can deduce from Equation (4.7) that the quantities  $u_j^{n+1}$  and  $v_j^{n+1}$  at time  $t^{n+1}$  are convex combinations of the quantities  $u_j^n$  and  $v_j^n$  at time  $t^n$  so long as we have  $\lambda_1 \geq 0$ ,  $\lambda_2 \geq 0$ ,  $\lambda_3 \geq 0$  and  $\lambda_4 \geq 0$ . We systematically have  $\lambda_3 \geq 0$  and  $\lambda_4 \geq 0$ . Condition  $\lambda_3 \geq 0$  leads to  $\Delta t \geq \varepsilon \Delta x / 2 = \Delta t_{min}$ . Condition  $\lambda_1 \geq 0$  breaks down to the study of the domain of positivity of a second-order polynomial in  $\Delta t$ , ultimately yielding  $\Delta t_{max}$  such as defined in (4.6).

Under these conditions, we can express  $u_j^{n+1}$  and  $v_j^{n+1}$  as convex combination of  $u_j^{n+1}$  and  $v_j^{n+1}$  and consequently we have the discrete maximum principle:

$$\min_{0 \leq j \leq N+1} (u_j^n, v_j^n) \leq \min_{1 \leq j \leq N} (u_j^{n+1}, v_j^{n+1}) \leq \max_{1 \leq j \leq N} (u_j^{n+1}, v_j^{n+1}) \leq \max_{0 \leq j \leq N+1} (u_j^n, v_j^n),$$

where we recall that indexes  $j = 0, N + 1$  stand for the ghost cells used to enforce boundary conditions. Under periodic boundary conditions (3.3) we have  $u_0^n = u_N^n$ ,  $u_{N+1}^n = u_1^n$ ,  $v_0^n = u_N^n$  and  $v_{N+1}^n = u_1^n$ . Meanwhile under hybrid boundary conditions (3.4), we have  $u_0^n = -v_1^n$ ,  $u_{N+1}^n = -v_N^n$ ,  $v_0^n = -u_0^n$  and  $v_{N+1}^n = -u_N^n$ . In both case (4.5) follows, which concludes the proof.  $\square$

A few remarks can be drawn from the previous studies:

- One observes that  $\Delta t_{max}$  is always strictly bigger than  $\Delta t_{min}$ :

$$(4.8) \quad \Delta t_{max} = \frac{\sigma \Delta x^2}{4} \frac{1 + \sqrt{1 + 2 \left(\frac{4\varepsilon}{\sigma \Delta x}\right)^2}}{2} > \frac{\sigma \Delta x^2}{4} \frac{\sqrt{2 \left(\frac{4\varepsilon}{\sigma \Delta x}\right)^2}}{2} = \sqrt{2} \Delta t_{min}.$$

It is always possible to make the ImEx1-ctr scheme stable irrespectively of  $\varepsilon$ .

- In the regime  $\varepsilon \ll \Delta x$  we have the classical parabolic stability condition:

$$(4.9) \quad \Delta t_{max} \underset{\varepsilon \rightarrow 0}{\sim} \frac{\sigma \Delta x^2}{4}.$$

In the regime  $\Delta x \ll \varepsilon$ , we have a classical hyperbolic stability conditions:

$$(4.10) \quad \Delta t_{max} \underset{\Delta x \rightarrow 0}{\sim} \frac{\varepsilon \Delta x}{\sqrt{2}}.$$

- The numerical diffusion in scheme (4.4), essential to stability, scales as  $\Delta t$ . This explains the need for a lower bound in  $\Delta t$  to avoid suppressing the stabilizing mechanism of the scheme. Possible ways of lifting this restriction are discussed at the end of this section.
- The  $l^\infty$ -stability is in general stronger than the  $l^2$ -stability. The convex combination form (4.7) ensures that the scheme does not produce any spurious numerical oscillations around discontinuities. This formulation also allows to retrieve the energy inequality (3.5).

**4.2.2.  $l^2$ -stability.** Under a slightly less limiting condition,  $l^2$ -stability follows.

**THEOREM 4.2.** *The ImEx1-ctr scheme (4.4) together with periodic boundary conditions (3.3) is  $l^2$ -diminishing under the condition:*

$$(4.11) \quad \Delta t \leq \Delta t_{max} = \frac{\sigma \Delta x^2}{4} \frac{1 + \sqrt{1 + \left(\frac{4\varepsilon}{\sigma \Delta x}\right)^2}}{2}.$$

*Proof.* Let us first notice that for all  $n \geq 0, 1 \leq j \leq N$ :

$$(E_j^n)^2 + (F_j^n)^2 = \frac{1}{4} \left( (u_j^n + v_j^n)^2 + (u_j^n - v_j^n)^2 \right) = \frac{1}{2} \left( (u_j^n)^2 + (v_j^n)^2 \right),$$

or in other words  $\|E_j^n\|_{l^2}^2 + \|F_j^n\|_{l^2}^2 = \frac{1}{2} (\|u_j^n\|_{l^2}^2 + \|v_j^n\|_{l^2}^2)$  so we can show that (3.5) holds for variables  $u$  and  $v$ . Because we assume periodic boundary conditions we only need to show that the Fourier modes  $\hat{u}_j^n = \hat{u} e^{i(n\omega \Delta t - jk \Delta x)}$  and  $\hat{v}_j^n = \hat{v} e^{i(n\omega \Delta t - jk \Delta x)}$ , with  $\omega, k \in \mathbb{R}$  are not amplified in  $l^2$ -norm by the scheme. We have:

$$w_j^{n+1} = A w_j^n, \quad w_j^n = \begin{pmatrix} \hat{u}_j^n \\ \hat{v}_j^n \end{pmatrix}, \quad A = \begin{pmatrix} 1 - a - ic - b & b \\ b & 1 - a + ic + b \end{pmatrix},$$

$$a = \frac{4M \Delta t^2}{\varepsilon^2 \Delta x^2} \sin^2 \left( \frac{k \Delta x}{2} \right), \quad c = \frac{M \Delta t}{\varepsilon \Delta x} \sin(k \Delta x), \quad b = \frac{\sigma M \Delta t}{2 \varepsilon^2}.$$

We can separate the hyperbolic structure from source and diffusion terms:

$$A = \frac{1}{2} (A_1 + A_2), \quad A_1 = \begin{pmatrix} 1 - 2a - i2c & 0 \\ 0 & 1 - 2a + i2c \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 - 2b & 2b \\ 2b & 1 - 2b \end{pmatrix}.$$

Matrix  $A_2$  is self-adjoint, it can be diagonalized in an orthogonal basis and consequently we have  $\|A_2\| = \rho(A_2) = |1 - 4b|$ . Because  $b$  is non-negative it is equivalent to  $-1 \leq 1 - 4b$  or  $2b \leq 1$ , i.e.  $M \frac{\sigma \Delta t}{\varepsilon^2} \leq 1$ , which always holds by definition of  $M$ .

The matrix  $A_1$  is diagonal so that we have  $\|A_1\|^2 = (1 - 2a)^2 + (2c)^2$ . Setting  $M' = 2M$  for convenience in our notations, we have:

$$\|A_1\|^2 = 1 - M' \frac{4\Delta t^2}{\varepsilon^2 \Delta x^2} \sin^2 \left( \frac{k\Delta x}{2} \right) \left( 2 - M' \left( 1 - \sin^2 \left( \frac{k\Delta x}{2} \right) \left( 1 - \frac{4\Delta t^2}{\varepsilon^2 \Delta x^2} \right) \right) \right).$$

Consequently in order to satisfy the condition  $\|A_1\|^2 \leq 1$  we simply need the condition:

$$2 - M' \left( 1 - \sin^2 \left( \frac{k\Delta x}{2} \right) \left( 1 - \frac{4\Delta t^2}{\varepsilon^2 \Delta x^2} \right) \right) \geq 0.$$

This condition holds if  $(4\Delta t^2)/(\varepsilon^2 \Delta x^2) \leq 1$ , so we consider the case  $(4\Delta t^2)/(\varepsilon^2 \Delta x^2) > 1$  in what follows. Using the definition of  $M$ , this last condition reduces to finding the positive root of a second-order polynomial equation, leading to (4.11), which concludes the proof.  $\square$

*Remark 4.3.* The first noticeable point here is that there is no  $\Delta t_{min}$  for the  $l^2$ -stability. The second is that even if the two maximum time steps (4.6) and (4.11) are equivalent in the regime  $\varepsilon \ll \Delta x$ , they differ by a factor  $\sqrt{2}$  in the regime  $\Delta x \ll \varepsilon$ . This is due to the fact the inequality  $\|A_2\| \leq 1$  saturates only in the regime  $\varepsilon \ll \Delta x$  and is far from optimal in the regime  $\Delta x \ll \varepsilon$ . The actual limit on the time step  $\Delta t_{max}$  for the  $l^2$ -stability is likely to be the same as that of  $l^\infty$ -stability.

**4.2.3. Accuracy.** Having shown stability, we can demonstrate that accuracy is maintained throughout every regime. It is called *uniformly asymptotic preserving*.

**THEOREM 4.4.** *Suppose that (4.11) holds. Then, the ImEx1-ctr scheme (4.4) provides  $c_j^n(E) = \mathcal{O}(\Delta x)$  and  $c_j^n(F) = \mathcal{O}(\Delta x)$  independently on  $\varepsilon$ .*

*Proof.* We make all necessary assumptions on the regularity of the solution  $(E, F)$  to (2.1) in this proof, and by an abuse of notation, we use the generic notation  $w_j^n = w(t^n, x_j)$ . The consistency error on  $E$  reads:

$$(4.12) \quad c_j^n(E) = \partial_t E_j^n + \frac{M}{\varepsilon} \partial_x F_j^n - \frac{M\Delta t}{\varepsilon^2} \partial_{xx} E_j^n + \mathcal{O}(\Delta t \|\partial_{tt} E\|_\infty) + \mathcal{O}\left(\frac{M\Delta x^2}{\varepsilon} \|\partial_{xxx} F\|_\infty\right) + \mathcal{O}\left(\frac{M\Delta t \Delta x^2}{\varepsilon^2} \|\partial_{xxxx} E\|_\infty\right).$$

By hypothesis,  $\|\partial_{xxx} F\|_\infty = \varepsilon \|\partial_{txx} E\| = \mathcal{O}(\varepsilon)$  and  $\Delta t = \mathcal{O}(\Delta x)$  and consequently the first two terms in Equation (4.12) are  $\mathcal{O}(\Delta x)$ . Furthermore,  $M\Delta t/\varepsilon^2 \leq 1/\sigma$  and consequently the last term is  $\mathcal{O}(\Delta x^2)$ . Factorizing by  $M$  and using (2.1b) provides:

$$\begin{aligned} c_j^n(E) &\underset{\mathcal{O}(\Delta x)}{=} M \left( \partial_t E_j^n + \frac{1}{\varepsilon} \left( \partial_x F_j^n - \frac{\Delta t}{\varepsilon} \partial_{xx} E_j^n - \frac{\sigma \Delta t}{\varepsilon} \partial_x F_j^n \right) \right) \\ &\underset{\mathcal{O}(\Delta x)}{=} M \left( \partial_t E_j^n - \partial_t E_j^{n+1} + \mathcal{O}\left(\frac{\Delta t^2}{\varepsilon}\right) \right) \underset{\mathcal{O}(\Delta x)}{=} \mathcal{O}\left(M\Delta t + \frac{M\Delta t^2}{\varepsilon}\right). \end{aligned}$$

where  $\underset{\mathcal{O}(\Delta x)}{=}$  denotes that we omitted the terms that have already been shown to be uniformly of order  $\mathcal{O}(\Delta x)$  in all regimes. Using analogous argument, we have  $M\Delta t^2/\varepsilon = \mathcal{O}(\varepsilon \Delta t/\sigma)$  so that the first order accuracy is achieved in all regimes.

The consistency on  $F$  reads:

$$(4.13) \quad c_j^n(F) = \partial_t F_j^n + \frac{M}{\varepsilon} \partial_x E_j^n - \frac{M\Delta t}{\varepsilon^2} \partial_{xx} F_j^n + M \frac{\sigma}{\varepsilon^2} F_j^n \\ + \mathcal{O}(\Delta t \|\partial_{tt} F\|_\infty) + \mathcal{O}\left(\frac{M\Delta x^2}{\varepsilon} \|\partial_{xxx} E\|_\infty\right) + \mathcal{O}\left(\frac{M\Delta t\Delta x^2}{\varepsilon^2} \|\partial_{xxxx} F\|_\infty\right).$$

The first and last terms of Equation (4.13) can be dealt with using the same approach as before. Regarding the middle term, we consider each regime: If  $\varepsilon \ll \Delta x$ , then  $\Delta t \approx \Delta x^2$  and  $M \approx \varepsilon^2/\Delta t$  which leads to  $M\Delta x^2/\varepsilon \approx \varepsilon \ll \Delta x$ . If  $\varepsilon \approx \Delta x$ , then  $M\Delta x^2/\varepsilon \leq \Delta x^2/\varepsilon \approx \Delta x$ . If  $\varepsilon \gg \Delta x$ , then  $\Delta x \leq \varepsilon$  and consequently  $M\Delta x^2/\varepsilon \leq \Delta x$ . Then the consistency error reads:

$$c_j^n(F) \underset{\mathcal{O}(\Delta x)}{=} M \left( \partial_t F_j^n + \frac{1}{\varepsilon} \left( \partial_x E_j^n - \frac{\Delta t}{\varepsilon} \partial_{xx} F_j^n \right) + \frac{\sigma}{\varepsilon^2} (F_j^n + \Delta t \partial_t F_j^n) \right) \\ \underset{\mathcal{O}(\Delta x)}{=} M \left( \partial_t F_j^n + \frac{1}{\varepsilon} \partial_x E_j^{n+1} + \frac{\sigma}{\varepsilon^2} F_j^{n+1} + \mathcal{O}\left(\frac{\Delta t^2}{\varepsilon} + \frac{\sigma \Delta t^2}{\varepsilon^2}\right) \right) \\ \underset{\mathcal{O}(\Delta x)}{=} \mathcal{O}\left(M\Delta t + \frac{M\Delta t^2}{\varepsilon} + \frac{\sigma M\Delta t^2}{\varepsilon^2}\right).$$

Using that  $\sigma M\Delta t^2/\varepsilon^2 \leq \Delta t$ , we retrieve order one accuracy in all regimes.  $\square$

*Remark 4.5.* In this proof, even though second order centered discretization of the fluxes have been used, the error term  $\mathcal{O}(M\Delta x^2/\varepsilon)$  degenerates to order 1 in the regime  $\Delta x \approx \varepsilon$ . This is not an issue since we only aim at first order accuracy. This aspect will be dealt with in Section 5 for the second order scheme.

**4.3. Alternative fluxes.** We have shown that scheme (4.4) possesses strong robustness property under conditions (4.6); however, the presence of a lower bound  $\Delta t_{min}$  to ensure that shock do not trigger spurious oscillations can be seen as too restrictive. It is possible to remove this minimal time step condition by choosing more diffusive approximations for the fluxes. We suggest the following:

$$(4.14a) \quad [\partial_x F]_j^{n+1} = F_{j+\frac{1}{2}}^{n+1} - F_{j-\frac{1}{2}}^{n+1}, \quad F_{j+\frac{1}{2}}^{n+1} = \frac{1}{2} \left( F_j^{n+1} + F_j^{n+1} + \lambda_{j+\frac{1}{2}} (E_{j+1}^n - E_j^n) \right),$$

$$(4.14b) \quad [\partial_x E]_j^{n+1} = E_{j+\frac{1}{2}}^{n+1} - E_{j-\frac{1}{2}}^{n+1}, \quad E_{j+\frac{1}{2}}^{n+1} = \frac{1}{2} \left( E_j^{n+1} + E_j^{n+1} + \lambda_{j+\frac{1}{2}} (F_{j+1}^n - F_j^n) \right),$$

where  $\lambda_{j+1/2} = \max\left(\frac{|F_{j+1}^n|}{E_{j+1}^n}, \frac{|F_j^n|}{E_j^n}\right)$ . This eventually leads to:

$$(4.14c) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M}{\varepsilon} \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} - \frac{M\Delta t}{\varepsilon^2} \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2} \\ - \frac{M\Delta x}{2\varepsilon} \frac{\lambda_{j+1/2} (E_{j+1}^n - E_j^n) - \lambda_{j-1/2} (E_j^n - E_{j-1}^n)}{\Delta x^2} = 0,$$

$$(4.14d) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{M}{\varepsilon} \frac{E_{j+1}^n - E_{j-1}^n}{2\Delta x} - \frac{M\Delta t}{\varepsilon^2} \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\Delta x^2} \\ - \frac{M\Delta x}{2\varepsilon} \frac{\lambda_{j+1/2} (F_{j+1}^n - F_j^n) - \lambda_{j-1/2} (F_j^n - F_{j-1}^n)}{\Delta x^2} = -M \frac{\sigma}{\varepsilon^2} F_j^n.$$

The diffusion term  $\frac{\lambda_{j+1/2} M \Delta x}{2} \partial_{xx} E$  is now controlled in all regimes since  $\lambda_{j+1/2}$  scales as  $|F| = \mathcal{O}(\varepsilon)$ . Furthermore, as  $F \sim \varepsilon \partial_x E / \sigma$ , the factor  $\lambda_{j+1/2}$  vanishes where the variation of  $E$  are small. In such areas, the scheme (4.14) effectively reduces to the ImEx1-ctr scheme (4.4). On the contrary in the vicinity of discontinuities this scheme reduces to an upwind-like scheme (see (B.1)). Rigorous proofs of these assertions are significantly more involved than what has been done previously and consequently they are only verified numerically.

**5. Second-order AP-schemes.** In line with the introduced formalism, the second order scheme has a general structure of the form:

$$(5.1a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_x F]_j^{n+1/2} = 0,$$

$$(5.1b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_x E]_j^{n+1/2} = -\frac{\sigma}{\varepsilon^2} F_j^{n+1/2}.$$

**5.1. Principle.** A naive approach would be to set:

$$(5.2a) \quad \frac{[\partial_x F]_j^{n+1/2} - [\partial_x F]_j^n}{\Delta t/2} = -\frac{1}{\varepsilon} [\partial_{xx} E]_j^n - \frac{\sigma}{\varepsilon^2} [\partial_x F]_j^{n+1/2},$$

$$(5.2b) \quad \frac{[\partial_x E]_j^{n+1/2} - [\partial_x E]_j^n}{\Delta t/2} = -\frac{1}{\varepsilon} [\partial_{xx} F]_j^n,$$

$$(5.2c) \quad F_j^{n+1/2} = \frac{F_j^{n+1} + F_j^n}{2},$$

leading to the scheme:

$$(5.3a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M_{1/2}}{\varepsilon} [\partial_x F]_j^n - \frac{M_{1/2} \Delta t}{\varepsilon} [\partial_{xx} E]_j^n = 0,$$

$$(5.3b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{M_{1/2}}{\varepsilon} [\partial_x E]_j^n - \frac{M \Delta t}{\varepsilon} [\partial_{xx} F]_j^n = -\frac{\sigma M_{1/2}}{\varepsilon^2} F_j^n,$$

with  $M_{1/2} = 1/(1 + \sigma \Delta t / (2\varepsilon^2))$ . Unfortunately, scheme (5.3) can be shown to have stability constraints in either  $l^2$  or  $l^\infty$  norms that scale as  $\Delta t_{max} = \mathcal{O}(\varepsilon \Delta t)$ . The idea is that our current choice of source term (5.2c) is not “implicit enough” and we have to shift it somehow a little bit more toward time  $t^{n+1}$ . We introduce a Reverse Runge-Kutta methodology<sup>1</sup> in order to treat the source term “as implicitly as possible”. A brief literature review and synthetic presentation of this approach for second order can be found in Appendix C. Going for second order with nice A-stability and L-stability properties, it boils down to reach time  $t^{n+1/2}$  in (5.2c) from time  $t^{n+1}$ :

$$(5.4) \quad \begin{aligned} F_j^{n+1/2} &= F_j^{n+1} - \frac{\Delta t}{2} [\partial_t F]_j^{n+1} \\ &= F_j^{n+1} + \frac{\Delta t}{2\varepsilon} \left( [\partial_x E]_j^{n+1} + \frac{\sigma}{\varepsilon} F_j^{n+1} \right) \\ &= \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right) F_j^{n+1} + \frac{\Delta t}{2\varepsilon} [\partial_x E]_j^n - \frac{\Delta t^2}{2\varepsilon^2} [\partial_{xx} F]_j^n. \end{aligned}$$

<sup>1</sup>Called *reflected* in [11] or *adjoint* in [21], it is investigated in [27, 36] (See Appendix C).

To the knowledge of the authors, this step differs essentially from other classes of ImEx schemes found in the literature [1, 7]. It is important for stability reasons that a similar backward step be used for the fluxes, starting from  $t^{n+1/2}$  and coming back to  $t^n$  because the term  $[\partial_{xx}E]_j^n$  cannot be brought to  $t^{n+1/2}$  without using a term of the form  $[\partial_{xxx}F]_j^n$ , which would increase the stencil:

$$\begin{aligned}
(5.5) \quad \frac{[\partial_x F]_j^{n+1/2} - [\partial_x F]_j^n}{\Delta t/2} &= -\frac{1}{\varepsilon} [\partial_{xx}E]_j^n - \frac{\sigma}{\varepsilon^2} \left( [\partial_x F]_j^{n+1/2} - \frac{\Delta t}{2} [\partial_{tx}F]_j^{n+1/2} \right) \\
&= -\frac{1}{\varepsilon} [\partial_{xx}E]_j^n - \frac{\sigma}{\varepsilon^2} \left( [\partial_x F]_j^{n+1/2} + \frac{\Delta t}{2\varepsilon} \left( [\partial_{xx}E]_j^n + \frac{\sigma}{\varepsilon} [\partial_x F]_j^{n+1/2} \right) \right) \\
&= -\frac{1}{\varepsilon} \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right) [\partial_{xx}E]_j^n - \frac{\sigma}{\varepsilon^2} \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right) [\partial_x F]_j^{n+1/2}.
\end{aligned}$$

For  $[\partial_x E]_j^{n+1/2}$ , we keep the formula (5.2b). Regarding the choice of discrete operator for the fluxes  $[\partial_x E]_j^n$ ,  $[\partial_x F]_j^n$  we have shown in Section 4 that the centered discretization is a valid choice and we use it as a reference in this section. Using these different elements we obtain the ImEx2-ctr second-order centered scheme:

$$\begin{aligned}
(5.6a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M_1}{\varepsilon} \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} - \frac{M_1^+ \Delta t}{2\varepsilon^2} \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2} &= 0, \\
(5.6b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{M_2}{\varepsilon} \frac{E_{j+1}^n - E_{j-1}^n}{2\Delta x} - \frac{M_2^+ \Delta t}{2\varepsilon^2} \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\Delta x^2} &= -\frac{\sigma M_2}{\varepsilon^2} \tilde{F}_j^n, \\
M_1 &= \frac{1}{1 + \frac{\sigma \Delta t}{2\varepsilon^2} \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right)}, \quad M_1^+ = \frac{1 + \frac{\sigma \Delta t}{2\varepsilon^2}}{1 + \frac{\sigma \Delta t}{2\varepsilon^2} \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right)}, \\
M_2 &= \frac{1 + \frac{\sigma \Delta t}{2\varepsilon^2}}{1 + \frac{\sigma \Delta t}{\varepsilon^2} \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right)}, \quad M_2^+ = \frac{1 + \frac{\sigma \Delta t}{\varepsilon^2}}{1 + \frac{\sigma \Delta t}{\varepsilon^2} \left( 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \right)}, \quad \tilde{F}_j^n = \frac{F_{j+1}^n + 4F_j^n + F_{j-1}^n}{6}.
\end{aligned}$$

The choice of  $\tilde{F}_j^n$  will be justified during the accuracy study. Meanwhile, we can see that factors  $M_1$ ,  $M_1^+$ ,  $M_2$  and  $M_2^+$  are essentially perturbations of factor  $M$ , allowing for a finer control. We emphasize that the second order method has the same overall structure as the first order one. In particular it is more compact in space than other ImEx approaches based on classical RK or LM methods [1, 8]. We now show that it is endowed with good stability and accuracy properties in all regimes.

**5.2.  $l^\infty$ -stability.** The ImEx2-ctr scheme (5.6) is no longer symmetric in  $E$  and  $F$ , we introduce the variables  $\tilde{u} = \sqrt{M_2}E + \sqrt{M_1}F$  and  $\tilde{v} = \sqrt{M_2}E - \sqrt{M_1}F$  that diagonalize the hyperbolic part of the scheme. We can see that in regimes where  $\Delta t \ll \varepsilon$ , we have  $\tilde{u} \rightarrow u$  and  $\tilde{v} \rightarrow v$ .

**THEOREM 5.1.** *The ImEx2-ctr scheme (5.6) together with periodic (3.3) or hybrid (3.4) boundary conditions is  $l^\infty$ -diminishing for the variables  $\tilde{u}, \tilde{v}$  provided that:*

$$(5.7) \quad \Delta t_{min} := \varepsilon \Delta x + \frac{\sigma \Delta x}{6} \leq \Delta t \leq \Delta t_{max} := \frac{\sigma \Delta x^2}{9/2} \frac{1 + \sqrt{1 + \frac{9}{2} \left( \frac{2\varepsilon}{\sigma \Delta x} \right)^2}}{2}.$$

*Proof.* The proof is similar to that of Theorem 4.1. Transposing equations (5.6) to the variables  $\tilde{u}$  and  $\tilde{v}$  and reordering the terms in the manner of Equation (4.7) we

obtain the stability conditions:

$$(5.8a) \quad 1 - \frac{M_+^+ \Delta t^2}{\varepsilon^2 \Delta x^2} - \frac{M_2 \sigma \Delta t}{3\varepsilon^2} \geq 0,$$

$$(5.8b) \quad \frac{M_+^+ \Delta t^2}{2\varepsilon^2 \Delta x^2} - \frac{\tilde{M} \Delta t}{2\varepsilon \Delta x} - \frac{M_2 \sigma \Delta t}{2 \times 6\varepsilon^2} \geq 0,$$

$$(5.8c) \quad \frac{M_2 \sigma \Delta t}{3\varepsilon^2} - \frac{M_-^+ \Delta t^2}{\varepsilon^2 \Delta x^2} \geq 0,$$

with  $\tilde{M} = \sqrt{M_1 M_2}$ ,  $M_\pm^+ = (M_1^+ \pm M_2^+)/2 > 0$ . If  $M_+^+ \Delta t^2 / (\varepsilon^2 \Delta x^2) \leq 2/3$  and  $M_2 \sigma \Delta t / (3\varepsilon^2) \leq 1/3$  then the first condition (5.8a) holds. By construction, the latter always holds. For the first, one computes:

$$(5.9) \quad M_1^+ = \frac{1 + \frac{\sigma \Delta t}{\varepsilon^2}}{\frac{3}{4} + \left(\frac{1}{2} + \frac{\sigma \Delta t}{2\varepsilon^2}\right)^2} \leq \frac{1 + \frac{\sigma \Delta t}{\varepsilon^2}}{\left(\frac{1}{2} + \frac{\sigma \Delta t}{2\varepsilon^2}\right)^2} \leq \frac{4}{1 + \frac{\sigma \Delta t}{\varepsilon^2}},$$

$$(5.10) \quad M_2^+ = \frac{1 + \frac{\sigma \Delta t}{\varepsilon^2}}{\frac{1}{2} + \left(\frac{1}{\sqrt{2}} + \frac{\sigma \Delta t}{\sqrt{2}\varepsilon^2}\right)^2} \leq \frac{1 + \frac{\sigma \Delta t}{\varepsilon^2}}{\left(\frac{1}{\sqrt{2}} + \frac{\sigma \Delta t}{\sqrt{2}\varepsilon^2}\right)^2} = \frac{2}{1 + \frac{\sigma \Delta t}{\varepsilon^2}}.$$

In the end we retrieve the bound  $M_+^+ = \frac{1}{2}(M_1^+ + M_2^+) \leq 3/(1 + (\sigma \Delta t)/(\varepsilon^2))$ . The study then reduces again to finding the root of a second-order polynomial equation and we find  $\Delta t_{max}$  as defined in (5.7). The second condition (5.8b) holds provided that:

$$\Delta t \geq \Delta t_{min} := \varepsilon \Delta x + \frac{\sigma \Delta x}{6} \geq \frac{\tilde{M}}{M_+^+} \varepsilon \Delta x + \frac{M_2}{M_+^+} \frac{\sigma \Delta x^2}{6},$$

since  $\tilde{M}/M_+^+ \leq 1$  and  $M_2/M_+^+ \leq 1$ . Finally the last condition (5.8c) is equivalent to:

$$\frac{3}{8} \frac{\frac{\sigma \Delta t}{\varepsilon^2}}{\left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)\right)} \frac{\Delta t^2}{\varepsilon^2 \Delta x^2} \leq 1.$$

This condition is verified if  $\Delta t \leq \frac{2}{3} \sigma \Delta x^2$  or if  $\Delta t \leq 2\sqrt{\frac{2}{3}} \varepsilon \Delta x$ . One of these two upper limits is always bigger than  $\Delta t_{max}$  so it is not restrictive.  $\square$

*Remark 5.2.* It is possible that the lower and upper bounds collide in the regimes  $\Delta x \approx \varepsilon$  and  $\Delta x \ll \varepsilon$ . This is due to the fact that the proposed  $\Delta t_{min}$  and  $\Delta t_{max}$  are not actual solutions of equations associated with conditions (5.8a) and (5.8b). However numerically solving these equations, we have found that in practice we always obtain  $\Delta t_{max} > \Delta t_{min}$  for the actual solutions of (5.8a) and (5.8b).

**5.3.  $l^2$ -stability.** We also have a lower-bound-free result for the  $l^2$ -norm.

**THEOREM 5.3.** *The ImEx1-ctr scheme (4.4) together with periodic boundary conditions (3.3) is  $l^2$ -diminishing for variables  $\tilde{u}$  and  $\tilde{v}$  under the condition:*

$$(5.11) \quad \Delta t \leq \Delta t_{max} = \frac{\sigma \Delta x^2}{6} \frac{1 + \sqrt{1 + 6 \left(\frac{2\varepsilon}{\sigma \Delta x^2}\right)^2}}{2}.$$

*Proof.* As we use periodic boundary conditions, we study again the Fourier modes  $\hat{u}_j^n = \hat{u} e^{i(n\omega \Delta t - jk \Delta x)}$  and  $\hat{v}_j^n = \hat{v} e^{i(n\omega \Delta t - jk \Delta x)}$ . The scheme writes  $w_j^{n+1} = Aw_j^n$ ,

with:

$$w_j^n = \begin{pmatrix} \hat{u}_j^n \\ \hat{v}_j^n \end{pmatrix}, \quad A = \begin{pmatrix} 1 - M_+^+ a - i\tilde{M}c - M_2b & M_2b - M_-^- a \\ M_2b - M_-^- a & 1 - M_+^+ a + i\tilde{M}c - M_2b \end{pmatrix},$$

$$a = \frac{2\Delta t^2}{\varepsilon^2 \Delta x^2} \sin^2 \left( \frac{k\Delta x}{2} \right), \quad c = \frac{\Delta t}{\varepsilon \Delta x} \sin(k\Delta x), \quad b = \frac{2 + \cos(k\Delta x)}{3} \frac{\sigma \Delta t}{2\varepsilon^2}.$$

We split the matrix  $A$  into  $A = \frac{1}{2}(A_1 + A_2)$  where:

$$A_1 = \begin{pmatrix} 1 - 2M_+^+ a - 2i\tilde{M}c & 0 \\ 0 & 1 - 2M_+^+ a + 2i\tilde{M}c \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 1 - 2M_2b & 2M_2b - 2M_-^- a \\ 2M_2b - 2M_-^- a & 1 - 2M_2b \end{pmatrix}.$$

Because  $A_1$  is a diagonal matrix we have  $\|A_1\|^2 = (1 - 4aM_+^+)^2 + (4\tilde{M})^2$ . Thus, setting  $M'_+ = 2M_+$  and  $\tilde{M}' = 2\tilde{M}$ , we obtain:

$$\|A_1\|^2 = 1 - M'_+ \frac{4\Delta t^2}{\varepsilon^2 \Delta x^2} \sin^2 \left( \frac{k\Delta x}{2} \right) \times$$

$$\left( 1 - M'_+ \left( \left( \frac{\tilde{M}'}{M'_+} \right)^2 - \sin^2 \left( \frac{k\Delta x}{2} \right) \left( \left( \frac{\tilde{M}'}{M'_+} \right)^2 - \frac{\Delta t^2}{\varepsilon^2 \Delta x^2} \right) \right) \right).$$

By construction, we have  $\tilde{M}' \leq M'_+$ . We want the following condition to hold:

$$1 - M'_+ \left( \left( \frac{\tilde{M}'}{M'_+} \right)^2 - \sin^2 \left( \frac{k\Delta x}{2} \right) \left( \left( \frac{\tilde{M}'}{M'_+} \right)^2 - \frac{\Delta t^2}{\varepsilon^2 \Delta x^2} \right) \right) \geq 0.$$

When  $\frac{\Delta t^2}{\varepsilon^2 \Delta x^2} \leq \frac{\tilde{M}'}{M'_+} \leq 1$  this condition is verified. Now we consider the case where  $\frac{\Delta t^2}{\varepsilon^2 \Delta x^2} \geq \frac{\tilde{M}'}{M'_+}$ . Then the condition holds if:

$$1 + \frac{\sigma \Delta t}{\varepsilon^2} \geq 6 \left( \left( \frac{\tilde{M}'}{M'_+} \right)^2 - \left( \left( \frac{\tilde{M}'}{M'_+} \right)^2 - \frac{\Delta t^2}{\varepsilon^2 \Delta x^2} \right) \right)$$

where we have used the fact that  $M'_+ = 2M_+^+$  and the inequality  $M_+^+ \leq 3/(1 + \sigma \Delta t/(\varepsilon^2))$ . In the end this leads to  $1 + \sigma \Delta t/(\varepsilon^2) \geq 6\Delta t^2/(\varepsilon^2 \Delta x^2)$ , which eventually leads to condition (5.11).  $\square$

**5.4. Accuracy.** The ImEx2-ctr is uniformly AP as shown in the following theorem. The proof is similar to that of Theorem 4.4 but it is more technical and is thus presented in Appendix A for the sake of legibility.

**THEOREM 5.4.** *Suppose that (5.7) and (5.11) holds. Then the ImEx2-ctr scheme (5.6) provides  $c_j^n(E) = O(\Delta x^2)$  and  $c_j^n(F) = O(\Delta x^2)$  independently of  $\varepsilon$ .*

**5.5. Alternative fluxes: MUSCL reconstruction.** The ImEx2-ctr scheme (5.6) has a restrictive stability condition in the regimes where  $\varepsilon \leq \Delta x$ . This means that the interval between upper and lower bounds on  $\Delta t$ , to ensure that no spurious



oscillations appear in the vicinity of discontinuities, almost collapses. However the use of a MUSCL method with customary approximation of the fluxes (Rusanov, HLL or Roe) suffers accuracy loss in the regime  $\varepsilon \approx \Delta x$  in a similar manner to what occurs with the first order scheme. We circumvent this and modify scheme (4.14) (which does not lose accuracy in this regime) with  $\lambda_{j+\frac{1}{2}} = 0$  in (4.14a) (no diffusion) and  $\lambda_{j+\frac{1}{2}} = 1$  with a MUSCL-like reconstruction (see for instance [38]) in (4.14b). Since the implicit flux  $F$  is re-injected in the  $E$  equation, the scheme eventually achieves second order accuracy in both unknowns. The slopes in this reconstruction are limited using a minmod limiter to preserve stability. This scheme is afterward referred to as ImEx2-minimod. Numerical evidence of the relevance of these choices are presented in Section 7.

**6. Euler-friction.** In this section, we adapt the schemes ImEx1-ctr (4.4) and ImEx2-ctr (5.6) to barotropic Euler-friction (2.7) that reads Equation (1.1) with:

$$(6.1) \quad \mathbf{w} = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}, \quad \mathbf{f}(\mathbf{w}) = \begin{pmatrix} \rho u \\ \rho u^2 + p(\rho) \end{pmatrix}, \quad \mathbf{S}(\mathbf{w}) = B\mathbf{w}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

with  $p'(\rho) > 0$ . We no longer assume that  $\sigma$  is constant, it may depend on space:  $\sigma(x)$ , leading to a steady solution for long times for which  $\rho$  can have arbitrary complex shape unlike in the  $\sigma$ -constant case where it must be linear. Furthermore, we assume that the flux is homogeneous, i.e.  $\mathbf{f}(\mathbf{w}) = A(\mathbf{w})\mathbf{w}$  where  $A(\mathbf{w})$  is the Jacobian of  $\mathbf{f}$ , and more specifically we consider  $p(\rho) = C\rho^\nu$ . The homogeneity of the flux, essentially depending on the choice of equation of state for the pressure term, is verified for a wide variety of models (see for instance [19]), including the full set of Euler equations.

**6.1. First-order AP-scheme.** We recall the full structure of the scheme:

$$(6.2) \quad \frac{\mathbf{w}_j^{n+1} - \mathbf{w}_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1} = -\frac{B}{\varepsilon^2} \sigma_j \mathbf{w}_j^{n+1}.$$

We need to design the term  $[\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1}$ . Assuming sufficient regularity, we have:

$$\partial_t \mathbf{f}(\mathbf{w}) = A(\mathbf{w}) \partial_t \mathbf{w} = -\frac{1}{\varepsilon} A(\mathbf{w}) \partial_x \mathbf{f}(\mathbf{w}) - \frac{\sigma}{\varepsilon^2} A(\mathbf{w}) B\mathbf{w}.$$

Formally, the identity  $[\mathbf{f}(\mathbf{w})]_j^{n+1} = [\mathbf{f}(\mathbf{w})]_j^n + \Delta t [\partial_t \mathbf{f}(\mathbf{w})]_j^n$  is rewritten:

$$(6.3) \quad [\mathbf{f}(\mathbf{w})]_j^{n+1} = [\mathbf{f}(\mathbf{w})]_j^n - \frac{\Delta t}{\varepsilon} [(A(\mathbf{w}) \partial_x \mathbf{f}(\mathbf{w}))]_j^n - \frac{\Delta t \sigma_j}{\varepsilon^2} [(A(\mathbf{w}) B\mathbf{w})]_j^*,$$

where the flux term is chosen explicit and the source term is chosen of the form:

$$(6.4) \quad [A(\mathbf{w}) B\mathbf{w}]_j^* = A(\mathbf{w}_j^n) B\mathbf{w}_j^*, \\ \mathbf{w}_j^* = \mathbf{w}_j^n - \frac{\Delta t}{\varepsilon} [\partial_x \mathbf{f}(\mathbf{w})]_j^n - \frac{\Delta t \sigma_j}{\varepsilon^2} B\mathbf{w}_j^* = I_{M_j} \left( \mathbf{w}_j^n - \frac{\Delta t}{\varepsilon} [\partial_x \mathbf{f}(\mathbf{w})]_j^n \right),$$

with  $I_{M_j} = \text{Diag}(1, M_j)$  and  $M_j = 1/(1 + \sigma_j \Delta t / \varepsilon^2)$  as in (4.3). This corresponds to fixing the non-linear part at time  $t^n$  explicit and the linear part at time  $t^{n+1}$  computed using (6.2) with explicit fluxes. All together, using that  $A(\mathbf{w})\mathbf{w} = \mathbf{f}(\mathbf{w})$  and applying (formally) the space derivative to (6.3) leads to:

$$(6.5a) \quad [\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1} = [\partial_x \mathbf{f}_M(\mathbf{w})]_j^n - \frac{\Delta t}{\varepsilon} [\partial_x (A_M(\mathbf{w}) \partial_x \mathbf{f}(\mathbf{w}))]_j^n,$$

$$(6.5b) \quad A_M(\mathbf{w}) = A(\mathbf{w}) I_M, \quad \mathbf{f}_M(\mathbf{w}) = A_M(\mathbf{w})\mathbf{w},$$

which, applied to (6.1), provides:

$$(6.5c) \quad \mathbf{f}_M(\mathbf{w}) = \begin{pmatrix} M\rho u \\ (2M-1)\rho u^2 + p(\rho) \end{pmatrix},$$

$$(6.5d) \quad A_M(\mathbf{w})\partial_x \mathbf{f}(\mathbf{w}) = \begin{pmatrix} M\partial_x(\rho u^2 + p(\rho)) \\ 2uM\partial_x(\rho u^2 + p(\rho)) + ((p'(\rho))^2 - u^2)\partial_x(\rho u) \end{pmatrix}.$$

Coming back to (6.2), the complete structure finally reads:

$$(6.6) \quad \frac{\mathbf{w}_j^{n+1} - \mathbf{w}_j^n}{\Delta t} + \frac{1}{\varepsilon} [I_M \partial_x \mathbf{f}_M(\mathbf{w})]_j^n - \frac{\Delta t}{\varepsilon^2} [I_M \partial_x (A_M \partial_x \mathbf{f}(\mathbf{w}))]_j^n = -\frac{\sigma_j M_j}{\varepsilon^2} B \mathbf{w}_j^n.$$

The choices made in this construction extends the ones made in the linear case in (4.2) to the non-linear Euler-Friction equation (6.1) while preserving a tractable formula. Indeed, using this generic construction to the linear HHE (2.1) yields exactly (4.2).

**6.1.1. Spatial discretization.** We must choose how to compute the terms  $[I_M \partial_x \mathbf{f}_M(\mathbf{w})]_j^n$  and  $[I_M \partial_x (A_M \partial_x \mathbf{f}(\mathbf{w}))]_j^n$ . For the fluxes we choose a modified Rusanov approach (see (4.14)) that reads:

$$(6.7) \quad [I_M \partial_x \mathbf{f}_M(\mathbf{w})]_j^n = I_{M_j} \left[ \frac{\mathbf{f}_{M_{j+1}}(\mathbf{w}_{j+1}^n) - \mathbf{f}_{M_{j-1}}(\mathbf{w}_{j-1}^n)}{2\Delta x} - \frac{\frac{\Delta x}{2} \lambda_{j+1/2} (\mathbf{w}_{j+1}^n - \mathbf{w}_j^n) - \lambda_{j-1/2} (\mathbf{w}_j^n - \mathbf{w}_{j-1}^n)}{\Delta x^2} \right],$$

where  $\lambda_{j+1/2} = \max(M_{j+1}|u_{j+1}|, M_j|u_j|)$ . Regarding the second-order derivatives in space, a classical centered discretization is used. In this discretization step, one must choose what value to use for  $M_{j+1/2}$  and  $u_{j+1/2}^n$ . Here the coefficient  $M_{j+1/2}$  is obtained by replacing  $\sigma_j$  by  $(\sigma_j + \sigma_{j+1})/2$  in  $M$  and the velocity is obtained with the approximation  $u_{j+1/2}^n = (\sqrt{\rho_{j+1}^n} u_{j+1}^n + \sqrt{\rho_j^n} u_j^n) / (\sqrt{\rho_{j+1}^n} + \sqrt{\rho_j^n})$ .

**6.1.2. Stability and accuracy.** The linearized version of the Euler-friction system is (up to a linear change of variable) equivalent to the HHE. Thus, the linear stability and truncation error analyses follow from that of the previous sections. Besides, we can obtain a positivity result on the mass density.

**THEOREM 6.1.** *The first-order ImEx scheme (6.6) with fluxes (6.7) for the Euler-friction equations preserves the positivity of the mass density under the condition:*

$$(6.8) \quad \Delta t_{max} \leq \frac{\sigma_{min} \Delta x^2}{2} \frac{\left(1 - \frac{\varepsilon |u|_{max}^n}{\sigma_{min} \Delta x}\right) + \sqrt{\left(1 - \frac{\varepsilon |u|_{max}^n}{\sigma_{min} \Delta x}\right)^2 + 4 \times \frac{2\varepsilon^2 \left(\left(|u|_{max}^n\right)^2 + c^2\right)}{\sigma_{max}^2 \Delta x^2}}}{2},$$

that is for  $n \geq 0$ ,  $(\rho_j^n \geq 0, \quad \forall j \in \{1, \dots, N\})$  implies  $(\rho_j^{n+1} \geq 0, \quad \forall j \in \{1, \dots, N\})$ .

*Proof.* Compute:

$$\begin{aligned} \rho_j^{n+1} &= \left(1 - \frac{\Delta t}{\varepsilon \Delta x} \frac{\lambda_{j+1/2} + \lambda_{j-1/2}}{2} - \frac{\Delta t^2}{\varepsilon^2 \Delta x^2} (M_{j+1/2} + M_{j-1/2}) \left((u_j^n)^2 + c^2\right)\right) \rho_j^n \\ &\quad + \frac{\Delta t}{\varepsilon \Delta x} \left(\frac{M_{j+1/2} \Delta t}{\varepsilon \Delta x} \left((u_{j+1}^n)^2 + c^2\right) + \frac{\lambda_{j+1/2}}{2} - \frac{M_{j+1} u_{j+1}^n}{2}\right) \rho_{j+1}^n \\ &\quad + \frac{M \Delta t}{\varepsilon \Delta x} \left(\frac{M_{j-1/2} \Delta t}{\varepsilon \Delta x} \left((u_{j-1}^n)^2 + c^2\right) + \frac{\lambda_{j-1/2}}{2} + \frac{M_{j-1} u_{j-1}^n}{2}\right) \rho_{j-1}^n. \end{aligned}$$

By construction  $\lambda_{j+1/2} \geq |M_{j+1}u_{j+1}|$  and  $\lambda_{j-1/2} \geq |M_{j+1}u_{j-1}|$ , thus the factor in front of  $\rho_{j\pm 1}^n$  are positive. The factor in front of  $\rho_j^n$  if the following inequality holds:

$$1 - \frac{\Delta t}{\varepsilon \Delta x} M_{max} |u|_{max}^n - \frac{2\Delta t^2}{\varepsilon^2 \Delta x^2} M_{max} \left( (|u|_{max}^n)^2 + c^2 \right) \geq 0,$$

where  $M_{max} = 1/(1 + \sigma_{min}\Delta t/\varepsilon^2)$ ,  $\sigma_{min} = \min_j \sigma_j$  and  $|u|_{max}^n = \max_j |u_j^n|$ . This leads to a second-order equation which positive solution is (6.8).  $\square$

**6.2. Second-order AP-scheme.** The second-order scheme is of the form:

$$(6.9) \quad \frac{\mathbf{w}_j^{n+1} - \mathbf{w}_j^n}{\Delta t} + \frac{1}{\varepsilon} [\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1/2} = -\frac{\sigma_j}{\varepsilon^2} B \mathbf{w}_j^{n+1/2},$$

where we need to define both  $[\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1/2}$  and  $\mathbf{w}^{n+1/2}$ . Concerning the source term, we use again the Reverse Runge-Kutta method as in (5.4):

$$(6.10) \quad \mathbf{w}_j^{n+1/2} = \mathbf{w}_j^{n+1} + \frac{\Delta t}{2\varepsilon} \left( [\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1} + \frac{\sigma_j}{\varepsilon} B \mathbf{w}_j^{n+1} \right).$$

The flux term  $[\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1}$  is given by (6.5). For the flux term  $[\partial_x \mathbf{f}(\mathbf{w})]_j^{n+1/2}$ , we rewrite (6.3) and (6.4) replacing  $\Delta t$  by  $\Delta t/2$  and the term  $\mathbf{w}_j^*$  at time  $t^{n+1/2}$  is computed by:

$$\begin{aligned} \mathbf{w}_j^* &= \mathbf{w}_j^n - \frac{\Delta t}{2\varepsilon} [\partial_x \mathbf{f}(\mathbf{w})]_j^n - \frac{\Delta t \sigma_j}{2\varepsilon^2} B \left( \mathbf{w}_j^* + \frac{\Delta t}{2\varepsilon} [\partial_x \mathbf{f}(\mathbf{w})]_j^n + \frac{\sigma_j \Delta t}{2\varepsilon^2} B \mathbf{w}_j^* \right) \\ &= I_{M_1} \mathbf{w}_j^n - \frac{\Delta t}{2\varepsilon} I_{M_1^+} [\partial_x \mathbf{f}(\mathbf{w})]_j^n, \end{aligned}$$

where  $M_1$  and  $M_1^+$  are defined in (5.6). Here, the source term is computed at time  $t^n$  using a Reverse Runge-Kutta step from time  $t^{n+1/2}$  (the \*-terms), as it is done in (5.5). All together, this leads to the scheme:

$$\begin{aligned} \frac{\mathbf{w}_j^{n+1} - \mathbf{w}_j^n}{\Delta t} + \frac{I_{M_3}}{\varepsilon} \left[ [\partial_x \mathbf{f}_{M_1}(\mathbf{w})]_j^n - \frac{\Delta t}{2\varepsilon} \left[ \partial_x \left( A_{M_1^+}(\mathbf{w}) \partial_x \mathbf{f}(\mathbf{w}) \right) \right]_j^n \right. \\ \left. - \frac{\sigma_j \Delta t}{2\varepsilon^2} B \left( [\partial_x \mathbf{f}_M(\mathbf{w})]_j^n - \frac{\Delta t}{\varepsilon} [\partial_x (A_M \partial_x \mathbf{f}(\mathbf{w}))]_j^n \right) \right] = -\frac{\sigma_j M_{3j}}{\varepsilon^2} B \mathbf{w}_j^n, \end{aligned}$$

where  $M_3 = 1/(1 + (\sigma \Delta t/\varepsilon^2)(1 + (\sigma \Delta t/(2\varepsilon^2))))$ . The two flux and diffusion terms are kept separated because the coefficient  $\sigma$  *a priori* depends on space and do not pass under the space derivative sign.

Regarding the spatial discretization, as in Section 5.5, we use a centered discretization of the diffusive terms, and Lax-Friedrichs-like fluxes (6.7) with  $\lambda = 0$  in  $\rho$  and  $\lambda = 1$  in  $\rho u$  combined with a MUSCL reconstruction on  $\rho u$  (in a manner similar to that of Subsection 5.5). Thus we have successfully extended the ImEx2-ctr scheme (5.6) to a nonlinear system with complex fluxes and non-constant  $\sigma$  coefficient. The uniform second order accuracy can be obtained in the manner of Theorem 5.4.

**7. Numerical results.** We consider three test-cases. The first one is a boundary-value problem of the HHE (2.1), for which an exact smooth solution can be derived. This analytical test-case is used to establish the order of the methods presented in this paper. The second test case is a Riemann problem used to demonstrate the capacity of the methods, in both linear and nonlinear settings, to handle correctly shocks. Finally the last test-case assesses how well the methods capture steady states in the nonlinear case with non-constant  $\sigma$ .

**7.1. Exact test-case.** We consider the HHE complemented with Dirichlet boundary conditions  $E_L$ ,  $E_R$ , imposed on the  $E$  variable respectively on the left and right boundaries of the domain. It can be shown that an exact solution of this problem is:

$$\begin{aligned} E(t, x) &= f(t)g(x) + \frac{E_R - E_L}{x_R - x_L}(x - x_L) + E_L, \\ F(t, x) &= \varepsilon f'(t)G(x) - \frac{\varepsilon}{\sigma}(E_R - E_L), \end{aligned}$$

for  $t \geq 0$ ,  $x \in [x_L, x_R]$ , where:

$$\begin{aligned} f(t) &= \alpha \frac{\lambda_+ e^{\lambda_- t} - \lambda_- e^{\lambda_+ t}}{\lambda_+ - \lambda_-} + \beta \frac{e^{\lambda_+ t} - e^{\lambda_- t}}{\lambda_+ - \lambda_-}, \quad \lambda_{\pm} = -\frac{\sigma}{2\varepsilon^2} \left( 1 \mp \sqrt{1 - \left( \frac{2\pi\varepsilon}{\sigma} \right)^2} \right), \\ g(x) &= \sin(\pi(x - x_L)), \quad G(x) = \frac{\cos(\pi(x - x_L))}{\pi}. \end{aligned}$$

The parameter  $\alpha$  determines the amplitude of the solution and we choose  $\beta = -\frac{\pi^2}{\sigma}\alpha$  in order for the initial condition to satisfy  $\partial_t f = \mathcal{O}(\varepsilon)$ . In the verification that follows we use  $\alpha = \sigma = 1$  for simplicity.

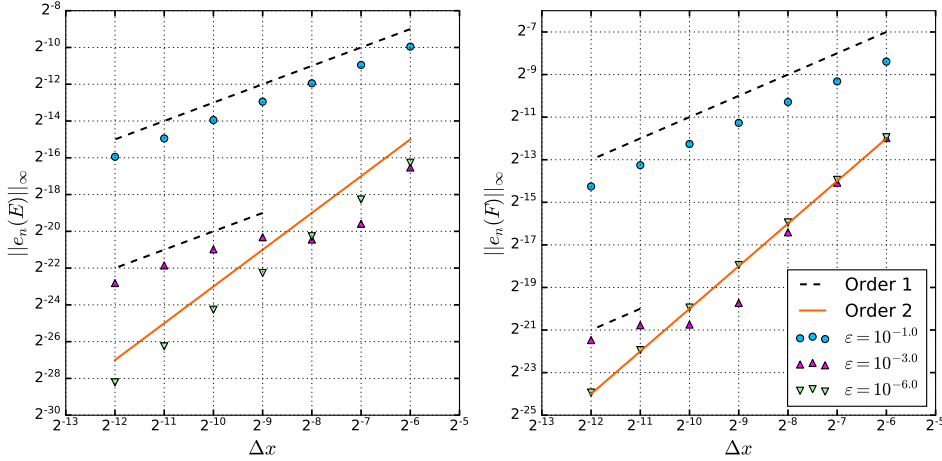


FIG. 1.  $l^\infty$ -norm of the global errors on  $E$  and  $F$  of the ImEx1-ctr scheme (4.4) compared to the analytical solution for different values of  $\varepsilon$  as a function of  $\Delta x$ .

In Figure 1, we can observe that the ImEx1-ctr scheme (4.4) is of order one when we consider mesh size such that  $\Delta x \leq \varepsilon$  and transition to a second order scheme in regimes where  $\varepsilon \ll \Delta x$ . For  $\varepsilon = 10^{-3}$  the chosen range of mesh sizes  $\Delta x$  encompasses both asymptotes with a transition around  $\Delta x \approx \varepsilon$ .

We observe a similar transition, in Figure 2, between two asymptotes related to the regimes  $\Delta x \leq \varepsilon$  and  $\Delta x \geq \varepsilon$ . These asymptotes are now of slope 2, which mean the ImEx2-ctr scheme (5.6) guarantees second order accuracy uniformly in all regimes.

Lastly, Figure 3 confirms that computing the flux on variable  $F$  using MUSCL reconstruction with a minmod limiter in the ImEx2-ctr scheme is still compatible with second order accuracy in all regimes.

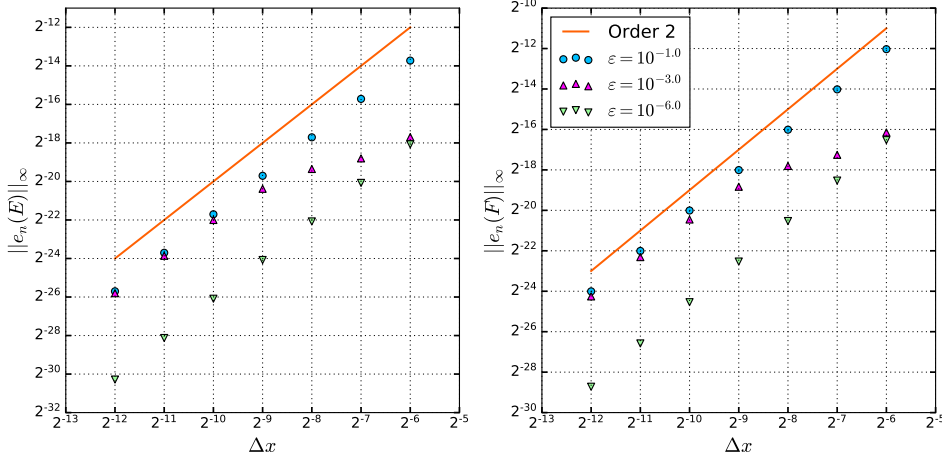


FIG. 2.  $l^\infty$ -norm of the global errors on  $E$  and  $F$  of the *ImEx2-ctr* scheme (5.6) compared to the analytical solution for different values of  $\varepsilon$  as a function of  $\Delta x$ .

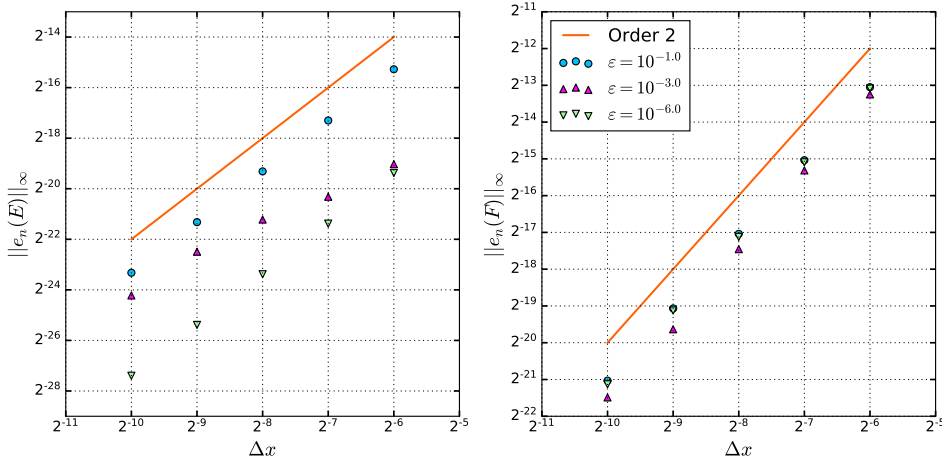


FIG. 3.  $l^\infty$ -norm of the global errors on  $E$  and  $F$  of the *ImEx2-minmod* scheme compared to the analytical solution for different values of  $\varepsilon$  as a function of  $\Delta x$ .

**7.2. Riemann problem.** The next test-case is that of an initial condition of the form of a Riemann problem. The linear case, which demonstrates the ability of the second-order method to be shock capturing without using slope limiters is presented in [Appendix D](#). For the more challenging case of the nonlinear Euler-friction equations (2.7), this test-case reads:

$$\rho(0, x) = \mathbf{1}_{\{x \leq \frac{x_R + x_L}{2}\}} \rho_L + \mathbf{1}_{\{x > \frac{x_R + x_L}{2}\}} \rho_R, \quad (\rho u)(0, x) = 0.$$

We present the results in [Figure 4](#) for the hyperbolic regime. The hyperbolic regime is more challenging than intermediate and parabolic regimes, for which the methods can function without limiters (see [Appendix D](#)). We can observe that the *ImEx2-ctr* scheme (5.6) triggers spurious oscillations whereas the *ImEx2-minmod* scheme essen-

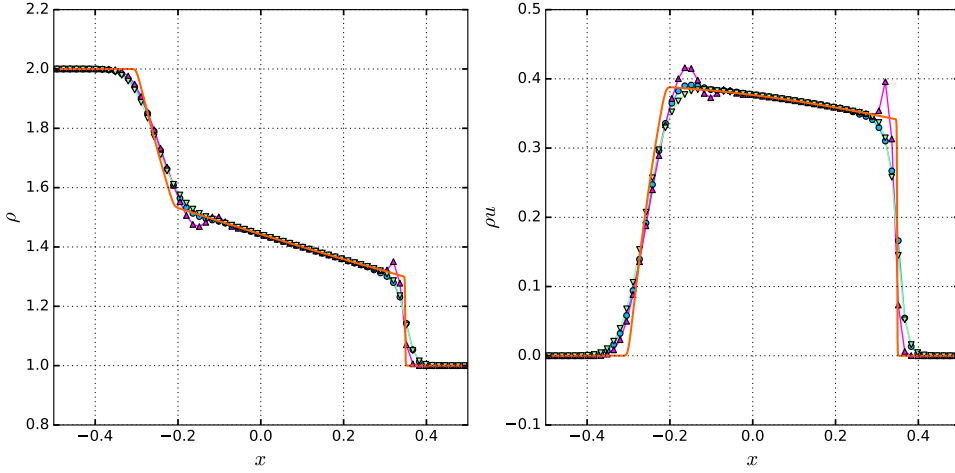


FIG. 4. Approximated solutions at time  $t = 0.15$ , with  $\sigma = 1$  and  $\varepsilon = 0.5$  (hyperbolic regime), computed respectively by the *ImEx2-ctr* method ( $\Delta$ ),  $\Delta t_{ctr} = 1.8\Delta t_{max}$ , from (5.11), the *ImEx2-minmod* method ( $\circ$ ),  $\Delta t_{minmod} = 1.18\Delta t_{max}$ , from (5.11) and the *MUSCL-Hancock* method with *Strang* splitting, *Reverse RK2* ( $\nabla$ ),  $\Delta t_{MH} = 0.9 \min(2\varepsilon^2/\sigma_{max}, \varepsilon\Delta x/(u_{max} + c))$ , using  $N = 64$  cells, with the latter scheme used as reference (solid line) with  $N = 2048$  cells.

tially shows equivalent performances as the classical MUSCL-Hancock scheme coupled with Strang splitting. It must be noted that in the nonlinear case the additional numerical diffusion of the *ImEx2-minmod* scheme does not remove completely the lower boundary condition  $\Delta t_{min}$ , however the interval of possible values for the time step is large enough to be compatible with all the velocities present in the problem, unlike the *ImEx2-ctr* scheme. As a result we have in the *ImEx2-minmod* scheme a method that incorporates enough numerical diffusion to be shock-capturing in the hyperbolic limit while retaining the asymptotic-preserving properties of the *ImEx2-ctr* in intermediate hyperbolic-parabolic and parabolic regimes.

**7.3. Steady state with non-constant relaxation coefficient.** The last test-case is concerned with the problem of correctly capturing steady states when the coefficient  $\sigma$  is non longer assumed spatially constant. This case is particularly interesting since it allows to consider both parabolic and hyperbolic regimes in one configuration. For a specific function  $\sigma(x)$ ,  $x \in [x_L, x_R]$ , the steady state solution can be easily characterized since we have for the isothermal Euler-friction equations:

$$\partial_x(\rho u) = 0, \quad \partial_x \left( \frac{(\rho u)^2}{\rho} + c^2 \rho \right) = -\frac{\sigma}{\varepsilon} \rho u, \quad \rho(t, x_L) = \rho_L, \quad \rho(t, x_R) = \rho_R,$$

so that if we set  $\rho u = \varepsilon a$  with  $a \in \mathbb{R}$  we obtain:

$$a = \frac{\rho_L \rho_L I_\sigma}{2\varepsilon^2 (\rho_L - \rho_R)} \left( \sqrt{1 + \frac{4c^2 \varepsilon^2 (\rho_L - \rho_R)^2}{\rho_L \rho_L I_\sigma^2}} - 1 \right), \quad I_\sigma = \int_{x_L}^{x_R} \sigma(x) dx,$$

$$\rho(x) = \frac{\bar{f}}{2c^2} \left( 1 + \sqrt{1 - \left( \frac{f_{min}}{\bar{f}} \right)^2} \right), \quad \bar{f} = \beta_\sigma f_R + (1 - \beta_\sigma) f_L, \quad f_{min} = 2\varepsilon a c$$

where  $f_R = \varepsilon^2 a^2 / \rho_R + c^2 \rho_R$ ,  $f_L = \varepsilon^2 a^2 / \rho_L + c^2 \rho_L$ ,  $\beta_\sigma = I_x / I_\sigma$  and  $I_x = \int_{x_L}^x \sigma(x) dx$ .

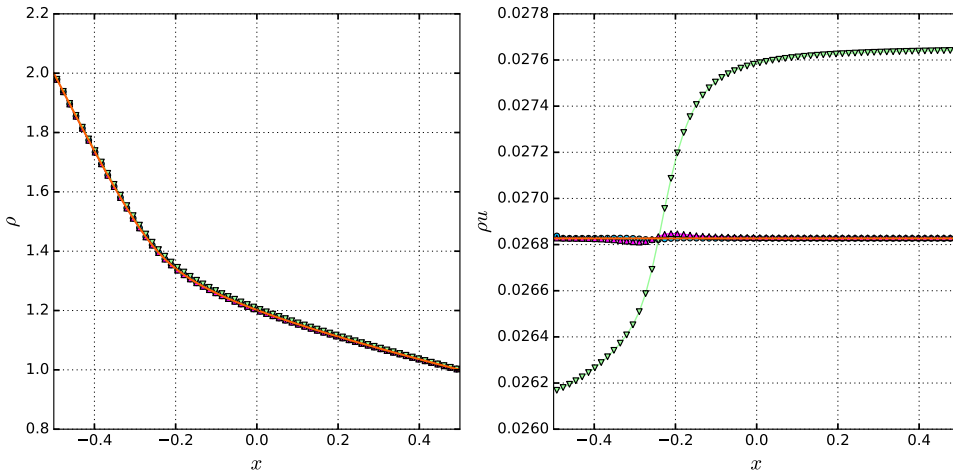


FIG. 5. Density  $\rho$  and momentum  $\rho u$  obtained at time  $t_f = 2$ , with non-constant  $\sigma$  and  $\varepsilon = 10^{-2}$ , respectively with the ImEx2-ctr method ( $\Delta$ ),  $\Delta t_{ctr} = 0.9 \Delta t_{max}$ , from (5.11), the ImEx2-minmod method ( $\circ$ ),  $\Delta t_{minmod} = 0.9 \Delta t_{max}$ , from (5.11) and the MUSCL-Hancock method with Strang splitting, Reverse RK2 ( $\nabla$ ),  $\Delta t_{MH} = 0.9 \min(2\varepsilon^2 / \sigma_{max}, \varepsilon \Delta x / (u_{max} + c))$ , using  $N = 64$  cells, with the steady state solution used as reference (solid line).

In Figure 5, we consider the case where:

$$\sigma(x) = (\sigma_{max} - \sigma_{min}) \frac{1 - \frac{2}{\pi} \arctan\left(\frac{x - \frac{x_c}{2}}{\tau}\right)}{1 - \frac{2}{\pi} \arctan\left(\frac{x_L - \frac{x_c}{2}}{\tau}\right)} + \sigma_{min},$$

where we define and chose  $\sigma_{max} = 1$ ,  $\sigma_{min} = 0.1$ ,  $x_c = (x_R + x_L) / 2$  and  $\tau = x_c / (2 \tan(\pi/2(1 - \gamma)))$ , so that  $1 - \pi/2 \arctan(x_c / (2\tau)) = \gamma$ , with  $\gamma = 0.1$  in order to obtain a  $C^\infty$ -function on  $[x_L, x_R]$ , going from  $\sigma_{max}$  at  $x_L$  to  $\sigma_{min}$  at  $x_R$  and reaches 90% of its right-value at  $x_c/2$ . What we observe is that although the ImEx2-ctr and ImEx2-minmod methods are not strictly well-balanced, in the sense of [9], especially where  $\sigma'$  is at its maximum value and at the boundaries for the minmod version, they still yield incomparably better results than the classical MUSCL-Hancock method coupled with Strang splitting.

**8. Conclusions.** We have derived two methods, ImEx2-ctr for the HHE and ImEx2-minmod for Euler-friction, which have the following properties:

- It is uniformly of second order accuracy in time and space in all regimes,
- Its stability condition does not collapse with  $\varepsilon$  in regimes where  $\varepsilon \rightarrow 0$ ,
- It can be readily extended to the nonlinear case of Euler-friction equations,
- In both linear and nonlinear setting it is shock-capturing; in the linear case and in the proper time step interval, ImEx2-ctr without slope limitation is even  $l^\infty$ -stable
- In both linear and nonlinear settings it captures accurately steady states when  $\sigma$  is not constant in space.

Let us underline the novelties of the methodology. First the methods are built relying on a coupled space-time discretization in the spirit of Lax-Wendroff approach,

but including an interesting implicit treatment of the sources and fluxes involving stiff relaxation. In particular, a specific integration in time of the source is introduced, from the class of mono-implicit methods, with nice stability properties which outmatch the classical second order implicit method [21, 27, 36]. This leads to fine evaluation and design of the stability properties, that are lost when time and space are decoupled as in most of the classic approaches. Besides, as opposed to the coupling of ODE solvers associated with WENO schemes for example, the proposed method is strongly compact in space and time in the spirit of a MUSCL-Hancock or Lax-Wendroff methods, while reaching interesting properties mentioned above.

This study is the building block of a series of natural extensions. We presently investigate ways to ensure the shock capturing property in a less diffusive way than ImEx2-minmod using slope limiters compatible with coupling between the fluxes and the source term, that is still maintaining the space-time approach. Besides, we are designing extensions of the schemes presented here to two-dimensional settings, to the problem of low-Mach limit and to more complex systems of equations such as that of the Euler-Poisson equations that can be used in plasmas physics. The ideas developed in this contribution are extended to plasma physics in a forthcoming paper and can also be used for the hyperbolic framework encountered in kinetic-fluid limits [34].

**Appendix A. Proof of Theorem 5.4.** We start with the consistency error on  $F$ . Isolating the spatial error term reads:

$$\begin{aligned} c_j^n(F) &= \partial_t F_j^{n+1/2} + \frac{M_2}{\varepsilon} \partial_x E_j^n - \frac{M_2^+ \Delta t}{2\varepsilon^2} \partial_{xx} F_j^n + \frac{\sigma M_2}{\varepsilon^2} F_j^n \\ &\quad + \mathcal{O}(\Delta t^2) + \mathcal{O}\left(\frac{M_2 \Delta x^2}{6\varepsilon} \left(\partial_{xxx} E_j^n + \frac{\sigma}{\varepsilon} \partial_{xx} F_j^n\right)\right) + \mathcal{O}\left(\frac{M_2^+ \Delta t}{2\varepsilon^2} \Delta x^2\right). \end{aligned}$$

Starting near the equilibrium manifold provides  $\frac{1}{\varepsilon} (\partial_{xxx} E_j^n + \frac{\sigma}{\varepsilon} \partial_{xx} F_j^n) = \mathcal{O}(1)$ . This legitimizes our choice of discretization for the source term. As  $M_2^+ \leq 2M$  (see (5.10)), these consistency error terms are controlled as in the proof of Theorem 4.4, via  $M\Delta t/\varepsilon^2 \leq 1$ . Now we must compute the time-error generated in the term:

$$\begin{aligned} &\frac{M_2}{\varepsilon} \partial_x E_j^n - \frac{M_2^+ \Delta t}{2\varepsilon^2} \partial_{xx} F_j^n \\ &= \frac{1}{\varepsilon \left(1 + \frac{\sigma \Delta t}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)\right)} \left( \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) \partial_x E_j^n - \frac{\Delta t}{2\varepsilon} \left(1 + \frac{\sigma \Delta t}{\varepsilon^2}\right) \partial_{xx} F_j^n \right) \\ &= \frac{1}{\varepsilon \left(1 + \frac{\sigma \Delta t}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)\right)} \left( \partial_x E_j^{n+1/2} + \mathcal{O}(\Delta t^2) + \frac{\sigma \Delta t}{2\varepsilon^2} \partial_x E_j^{n+1} + \mathcal{O}\left(\frac{\Delta t^3}{\varepsilon^2}\right) \right). \end{aligned}$$

This inequality produces two new error terms that must be controlled. Noticing that  $x \mapsto x^2/(1+x(1+x/2))$  is strictly increasing on  $\mathbb{R}_+$  and using the hypothesis  $\Delta t = \mathcal{O}\left(\Delta x \left(\sigma \Delta x + \sqrt{(\sigma \Delta x)^2 + \varepsilon^2}\right)\right)$ , then:

$$\begin{aligned} \frac{\Delta t^2}{\varepsilon \left(1 + \frac{\sigma \Delta t}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)\right)} &= \frac{\varepsilon^3 \Delta t^2}{\varepsilon^2 (\varepsilon^2 + \sigma \Delta t) + (\sigma \Delta t)^2 / 2} \\ &= \mathcal{O} \left( \varepsilon \Delta x^2 \frac{\varepsilon^2 \left(\sigma \Delta x + \sqrt{(\sigma \Delta x)^2 + \varepsilon^2}\right)^2}{\frac{\varepsilon^2}{2} \left(\sigma \Delta x + \sqrt{(\sigma \Delta x)^2 + \varepsilon^2}\right)^2 + \frac{\varepsilon^4}{2} + (\varepsilon \Delta t)^2 / 2} \right) = \mathcal{O}(\varepsilon \Delta x^2). \end{aligned}$$



Similarly,  $x \mapsto x^3 / (1 + x(1 + x/2))$  is strictly increasing on  $\mathbb{R}_+$  and using that  $\Delta t = \mathcal{O}\left(\Delta x \sqrt{(\sigma \Delta x)^2 + \varepsilon^2}\right)$ , then:

$$\begin{aligned} \frac{\Delta t^3}{\varepsilon^3 \left(1 + \frac{\sigma \Delta t}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)\right)} &= \mathcal{O}\left(\varepsilon \Delta x^2 \frac{\Delta x (\Delta x^2 + \varepsilon^2) \sqrt{\Delta x^2 + \varepsilon^2}}{\varepsilon^4 + \varepsilon^2 \Delta x \sqrt{\Delta x^2 + \varepsilon^2} + \Delta x^2 (\Delta x^2 + \varepsilon^2)}\right) \\ &= \mathcal{O}(\varepsilon \Delta x^2), \end{aligned}$$

using  $\sqrt{\Delta x^2 + \varepsilon^2} \leq \Delta x$  in the numerator and observing that the denominator is smaller than  $\Delta x^2 (\Delta x^2 + \varepsilon^2)$ . Coming back to  $c_j^n(F)$ , this provides:

$$\begin{aligned} c_j^n(F) &= \frac{C_j^N(F)}{1 + \frac{\sigma \Delta t}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)} + \mathcal{O}(\Delta x^2), \\ C_j^N(F) &= \left(1 + \frac{\sigma \Delta t}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right)\right) \partial_t F_j^{n+1/2} \\ &\quad + \frac{1}{\varepsilon} \left(\partial_x E_j^{n+1/2} + \frac{\sigma \Delta t}{2\varepsilon^2} \partial_x E_j^{n+1}\right) + \frac{\sigma}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) F_j^n \end{aligned}$$

Using the identities  $\partial_t F_j^{n+1/2} + (1/\varepsilon) \partial_x E_j^{n+1/2} = -(\sigma/\varepsilon^2) F_j^{n+1/2}$  and  $F_j^n + \Delta t \partial_t F_j^{n+1/2} = F_j^{n+1} + \mathcal{O}(\Delta t^3)$  leads to:

$$C_j^N(F) = \frac{\sigma}{\varepsilon^2} \left( \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) F_j^{n+1} + \frac{\Delta t}{2\varepsilon} \partial_x E_j^{n+1} - F_j^{n+1/2} + \mathcal{O}\left(\left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) \Delta t^3\right) \right).$$

Now, using  $\partial_x E_j^{n+1} + (\sigma/\varepsilon) F_j^{n+1} = \partial_t F_j^{n+1}$  provides

$$\begin{aligned} C_j^N(F) &= \frac{\sigma}{\varepsilon^2} \left( F_j^{n+1} - F_j^{n+1/2} - \frac{\Delta t}{2} \partial_t F_j^{n+1} + \mathcal{O}\left(\left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) \Delta t^3\right) \right) \\ &= \mathcal{O}\left(\frac{\sigma \Delta t^2}{\varepsilon^2}\right) + \mathcal{O}\left(\frac{\sigma}{\varepsilon^2} \left(1 + \frac{\sigma \Delta t}{2\varepsilon^2}\right) \Delta t^3\right) = \mathcal{O}(\Delta x^2) \end{aligned}$$

using the previous estimates.

Concerning  $E$ , we have:

$$c_j^n(E) = \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M_1}{\varepsilon} \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} - \frac{M_1^+ \Delta t}{2\varepsilon^2} \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2}.$$

First, the flux discretization provides:

$$\frac{M_1}{\varepsilon} \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} = \frac{M_1}{\varepsilon} \partial_x F_j^n + \frac{M_1 \Delta x^2}{2\varepsilon} \partial_{xxx} F_j^n + \mathcal{O}\left(\frac{M_1}{\varepsilon} \Delta x^4\right),$$

with  $\partial_{xxx} F_j^n = \mathcal{O}(\varepsilon)$  again such that the error is  $\mathcal{O}(\Delta x^2)$ . Second, the diffusion term provides:

$$\frac{M_1^+ \Delta t}{2\varepsilon^2} \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2} = \frac{M_1^+ \Delta t}{2\varepsilon^2} \partial_{xx} E_j^n + \mathcal{O}\left(\frac{M_1^+ \Delta t}{2\varepsilon^2} \Delta x^2\right),$$

but  $M_1^+ \Delta t / (2\varepsilon^2) \leq 1/\sigma = \mathcal{O}(1)$  so that  $\mathcal{O}(M_1^+ \Delta t \Delta x^2 / (2\varepsilon^2)) = \mathcal{O}(\Delta x^2)$ . Under the usual stability conditions, we have  $\mathcal{O}(\Delta t^2) = \mathcal{O}(\Delta x^2)$ . Now we have:

$$c_j^n(E) = \partial_t E_j^{n+1/2} + \frac{M_1}{\varepsilon} \partial_x F_j^n - \frac{M_1^+ \Delta t}{2\varepsilon^2} \partial_{xx} E_j^n + \mathcal{O}(\Delta x^2),$$

which leads to:

$$\begin{aligned} c_j^n(E) &= \frac{C_j^n(E)}{1 + \frac{\sigma\Delta t}{2\varepsilon^2} \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2}\right)} + \mathcal{O}(\Delta x^2) \\ C_j^n(E) &= \left( \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2} \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2}\right)\right) \partial_t E_j^{n+1/2} + \frac{1}{\varepsilon} \partial_x F_j^n - \frac{\Delta t}{2\varepsilon^2} \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2}\right) \partial_{xx} E_j^n \right) \\ &= \frac{-\Delta t}{2\varepsilon} \left( \frac{\partial_x F_j^{n+1/2} - \partial_x F_j^n}{\Delta t/2} + \frac{1}{\varepsilon} \partial_{xx} E_j^n \right. \\ &\quad \left. + \frac{\sigma}{\varepsilon^2} \left( \partial_x F_j^{n+1/2} + \frac{\Delta t}{2\varepsilon} \left( \partial_{xx} E_j^n + \frac{\sigma}{\varepsilon} \partial_x F_j^{n+1/2} \right) \right) \right). \end{aligned}$$

Using the usual Taylor expansions we obtain in the end:

$$c_j^n(E) = \mathcal{O} \left( \frac{\Delta t^2}{\varepsilon \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2}\right)} \right) + \mathcal{O} \left( \frac{\Delta t^3}{\varepsilon^4 \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2}\right) \left(1 + \frac{\sigma\Delta t}{2\varepsilon^2}\right)} \right) + \mathcal{O}(\Delta x^2),$$

and the estimations found for  $c_j^n(F)$  leads to simplify  $c_j^n(E) = \mathcal{O}(\Delta x^2)$ .

**Appendix B. Alternative discretizations for order 1 ImEx schemes: upwind discretization.** Although we have shown that scheme (4.4) demonstrates strong robustness property under the right set of stability conditions (4.6), the presence of a lower bound to ensure that shock do not trigger spurious oscillations can be seen as too restrictive. An attempt to remove this lower bound constraint can be made in using an upwind discretization of the fluxes, leading to the ImEx1-upwd scheme:

$$(B.1a) \quad \frac{E_j^{n+1} - E_j^n}{\Delta t} + \frac{M}{\varepsilon} \frac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} - \left( \frac{M\Delta x}{2\varepsilon} + \frac{M\Delta t}{\varepsilon^2} \right) \frac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\Delta x^2} = 0,$$

$$(B.1b) \quad \frac{F_j^{n+1} - F_j^n}{\Delta t} + \frac{M}{\varepsilon} \frac{E_{j+1}^n - E_{j-1}^n}{2\Delta x} - \left( \frac{M\Delta x}{2\varepsilon} + \frac{M\Delta t}{\varepsilon^2} \right) \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\Delta x^2} = -M \frac{\sigma}{\varepsilon^2} F_j^n.$$

**THEOREM B.1.** *Assuming that periodic (3.3) or hybrid (3.4) boundary conditions are used, the ImEx1-upwd scheme (B.1) is  $l^\infty$ -diminishing for variables  $u, v$  under the condition:*

$$(B.2) \quad \Delta t \leq \Delta t_{max} := \frac{\sigma\Delta x^2}{4} \left( \left( \frac{1}{2} - \frac{\varepsilon}{\sigma\Delta x} \right) + \sqrt{\left( \frac{\varepsilon}{\sigma\Delta x} - \frac{1}{2} \right)^2 + 2 \left( \frac{2\varepsilon}{\sigma\Delta x} \right)^2} \right).$$

*Proof.* The proof is very similar to that of [Theorem 4.1](#), using a form of the scheme (B.1) equivalent to the reordering (4.7) there is no longer any lower bound condition and the only stability restriction is

$$1 - \frac{2M\Delta t^2}{\varepsilon^2\Delta x^2} - \frac{M\Delta t}{\varepsilon\Delta x} - \frac{M\sigma\Delta t}{2\varepsilon^2} \geq 0,$$

which leads to (B.2) □

**THEOREM B.2.** *In the regime  $\Delta x \approx \varepsilon$ , the ImEx1-upwd scheme (B.1) loses accuracy in the sense  $c_j^n(E) = \mathcal{O}(1)$ .*

*Proof.* As compared to the consistency error conducted on scheme (4.4) we have an additional error term for variable  $E$  of the form  $(M\Delta x/2\varepsilon)\partial_{xx}E$ . However in the regime  $\varepsilon \approx \Delta x$ , we have  $M \approx 1$  and  $\Delta x/\varepsilon \approx 1$  and consequently this new term is of order  $\mathcal{O}(\Delta x^0)$  in this regime.  $\square$

*Remark B.3.* The impossibility to use the classical upwind discretization is not commonly discussed in literature but here we have shown that it does not yield an asymptotic preserving scheme.

### Appendix C. Reverse Runge-Kutta framework.

The approach we have used in order to integrate in time the source term is essential in the construction of the time-space-ImEx method proposed in the paper. We have called it a Reverse Runge-Kutta method. It has been introduced in the classical monographs [12], where it is referred to as *reflected*, and [22], under the denomination *adjoint* from a classical explicit RK method. Even if some general properties of these this specific class of method are provided in these works, the efficiency and computational aspects were rather investigated in [39, 14] (where they are called implicit endpoint quadrature formulas and mono-implicite RK Formulae respectively), and more recently in [28] and [37] (called inverse RK schemes and implicit RK methods obtained as a result of the inversion of explicit methods).

Before introducing the core principle of the approach, since it is essentially based on explicit Runge-Kutta methods, we introduce first-order and second order RK methods and the related notations and Butcher arrays. The general form of the autonomous ODE we want to solve reads:

$$d_t U = f(U),$$

where  $f$  is assumed to verify the assumptions of the Cauchy-Lipschitz theorem. With these notations the explicit Euler method and its corresponding Butcher array read:

$$\begin{array}{c|c} 0 & \\ \hline & 1 \end{array} \quad U_{n+1} = U_n + \Delta t f(U_n).$$

The general form of explicit second-order methods parametrized by  $\alpha \in (0, 1]$  reads:

$$\begin{array}{c|cc} 0 & & \\ \alpha & \alpha & \\ \hline & 1 - \frac{1}{2\alpha} & \frac{1}{2\alpha} \end{array} \quad \begin{array}{l} U_{n_1} = U_n + \alpha \Delta t f(U_n), \\ U_{n+1} = U_n + \Delta t \left(1 - \frac{1}{2\alpha}\right) f(U_n) + \frac{\Delta t}{2\alpha} f(U_{n_1}). \end{array}$$

For  $\alpha = 1$  we retrieve the Heun method and for  $\alpha = 1/2$ , the classical RK2:

$$\begin{aligned} U_{n_1} &= U_n + \frac{\Delta t}{2} f(U_n), \\ U_{n+1} &= U_n + \Delta t f(U_{n_1}). \end{aligned}$$

The principle of the Reverse Runge-Kutta method is the following: we substitute  $-t$  to  $t$  (reversed time) and we apply the explicit method starting from the final point. For instance, at order one and using the explicit Euler method we obtain:

$$U_n = U_{n+1} - \Delta t f(U_{n+1}),$$

which boils down to the backward Euler method:

$$U_{n+1} = U_n + \Delta t f(U_{n+1}).$$

At second order, we obtain:

$$U_n = U_{n+1} - \Delta t \left(1 - \frac{1}{2\alpha}\right) f(U_{n+1}) - \frac{\Delta t}{2\alpha} f(U_{n+1} - \alpha \Delta t f(U_{n+1})),$$

which can be rewritten as:

$$U_{n+1} = U_n + \Delta t \left(1 - \frac{1}{2\alpha}\right) f(U_{n+1}) + \frac{\Delta t}{2\alpha} f(U_{n+1} - \alpha \Delta t f(U_{n+1})).$$

Regarding absolute stability, the general stability function of an explicit second-order method is

$$R(z) = 1 + z + \frac{z^2}{2}.$$

Consequently, the stability function  $R^*(z)$  of second order Reverse RK methods can be obtained by substituting  $-z \rightarrow z$  and taking the inverse of  $R$ , leading to:

$$R^*(z) = \frac{1}{R(-z)}.$$

Because  $1/|1 - z + z^2/2| \leq 1$  for all  $z \in \mathbb{C}^-$  the second-order Reverse RK methods are  $A$ -stable and even  $L$ -stable.

These Reverse RK methods can be recast within the framework of classical RK methods (the Butcher array can be obtained explicitly from the original one [22]):

$$\begin{cases} U_{n_1} = U_n + \frac{\Delta t}{2\alpha} f(U_{n_1}) + \Delta t \left(1 - \frac{1}{2\alpha}\right) f(U_{n_2}) - \alpha \Delta t f(U_{n_2}), \\ U_{n_2} = U_n + \frac{\Delta t}{2\alpha} f(U_{n_1}) + \Delta t \left(1 - \frac{1}{2\alpha}\right) f(U_{n_2}), \\ U_{n+1} = U_n + \frac{\Delta t}{2\alpha} f(U_{n_1}) + \Delta t \left(1 - \frac{1}{2\alpha}\right) f(U_{n_2}). \end{cases}$$

One notices that  $U_{n_1} = U_{n_2} - \alpha \Delta t f(U_{n_2})$ , so that it suffices to solve for  $U_{n_2}$  directly. The corresponding Butcher arrays for  $\alpha \in (0, 1]$  and  $\alpha = 1/2$  (Reverse RK2 method used in the paper) read:

$$\begin{array}{c|cc} 1 - \alpha & \frac{1}{2\alpha} & 1 - \frac{1}{2\alpha} - \alpha \\ & \frac{1}{2\alpha} & 1 - \frac{1}{2\alpha} \\ \hline & \frac{1}{2\alpha} & 1 - \frac{1}{2\alpha} \end{array} \quad \begin{array}{c|cc} \frac{1}{2} & 1 & -\frac{1}{2} \\ & 1 & 0 \\ \hline & 1 & 0 \end{array}$$

which shows that the Reverse RK methods are stiffly accurate and have excellent stability properties. They can be shown to be computationally efficient as mono-implicit methods.

**Appendix D. Riemann problem, second order and  $l^\infty$ -stability without limitation.** We consider in this section the same Riemann-problem initial condition as in [Subsection 7.2](#), which reads in the linear case:

$$E(0, x) = \mathbf{1}_{\{x \leq \frac{x_R + x_L}{2}\}} E_L + \mathbf{1}_{\{x > \frac{x_R + x_L}{2}\}} E_R, \quad F(0, x) = 0,$$

and in the nonlinear case

$$\rho(0, x) = \mathbf{1}_{\{x \leq \frac{x_R + x_L}{2}\}} \rho_L + \mathbf{1}_{\{x > \frac{x_R + x_L}{2}\}} \rho_R, \quad (\rho u)(0, x) = 0.$$

We show that for both the linear case in the hyperbolic regime (meaning  $\varepsilon = 0.5$ ) and the nonlinear case in the intermediate parabolic-hyperbolic regime (meaning  $\varepsilon = 5.10^{-2}$  and  $\varepsilon = 10^{-2}$ ), the second-order centered methods do not trigger spurious oscillations around discontinuities without using additional numerical diffusion or slope limiters, provided that the time step falls within the acceptable range of values determined in [Equation \(5.7\)](#) implying  $l^\infty$ -stability. Within this regimes, the methods are quite accurate.

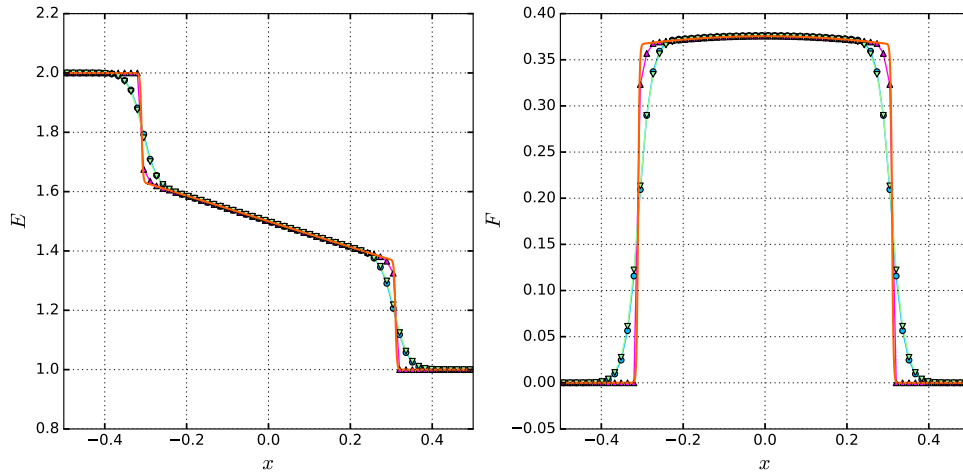


FIG. 6. Variables  $E$  (left) and  $F$  (right) obtained at time  $t_f = 20\Delta t_{ctr} \approx 0.15$ , for  $\varepsilon = 0.5$ ,  $E_L = 2$  and  $E_R = 1$ , respectively with the *ImEx2-ctr* method ( $\Delta$ ),  $\Delta t_{ctr} = (\Delta t_{max} + \Delta t_{min})/2$ , from [\(5.8a\)](#) and [\(5.8b\)](#), *ImEx2-minmod* method ( $\circ$ ),  $\Delta t_{minmod} = 1.2\Delta t_{max}$ , from [\(5.11\)](#) and *MUSCL-Hancock* method with *Strang* splitting, *Reverse RK2* ( $\nabla$ ),  $\Delta t_{MH} = 0.9 \min(2\varepsilon^2/\sigma_{max}, \varepsilon\Delta x/(u_{max} + c))$ , using  $N = 64$  cells, with the latter scheme used as reference (solid line) with  $N = 2048$  cells.

For the linear hyperbolic regime case, shown in [Figure 6](#), one observes that the *ImEx2-ctr* method does not generate oscillations in the vicinity of the discontinuity but, as expected, it is also significantly less diffusive than the *ImEx2-minmod* and *MUSCL-Hancock* methods, which both achieve similar performances. We emphasize however that in this case the estimations for  $\Delta t_{min}$  and  $\Delta t_{max}$  provided by [\(5.7\)](#) collide and therefore one must compute numerically the bound  $\Delta t_{min}$  and  $\Delta t_{max}$  corresponding to conditions [\(5.8a\)](#) and [\(5.8b\)](#) and choose  $\Delta t$  within the proper interval  $[\Delta t_{min}, \Delta t_{max}]$ .

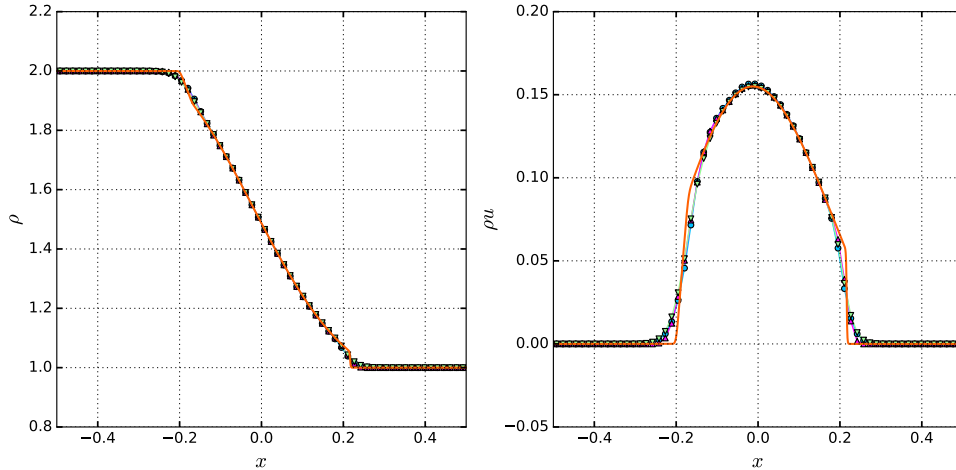


FIG. 7. Mass  $\rho$  and momentum  $\rho u$  obtained at time  $t_f = 0.01$  for  $\varepsilon = 5 \times 10^{-2}$ ,  $\rho_L = 2$  and  $\rho_R = 1$ , respectively with the *ImEx2-ctr* method ( $\triangle$ ),  $\Delta t_{ctr} = 1.87\Delta t_{max}$ , from (5.11), *ImEx2-minmod* method ( $\circ$ ),  $\Delta t_{minmod} = 0.9\Delta t_{max}$ , from (5.11) and *MUSCL-Hancock* method with Strang splitting, Reverse *RK2* ( $\nabla$ ),  $\Delta t_{MH} = 0.9 \min(2\varepsilon^2/\sigma_{max}, \varepsilon\Delta x/(u_{max} + c))$ , using  $N = 64$  cells, with the latter scheme used as reference (solid line) with  $N = 2048$  cells.

It is remarkable to observe that unlike in the case of purely hyperbolic equations, for which second order methods always require slope limiters in order to be stable around shocks, the presence of a source term guarantees that there is always an interval of possible time steps for which the *ImEx2-ctr* method can be  $l^\infty$ -stable.

Regarding the nonlinear intermediate and diffusive regimes case ( $\varepsilon = 5 \times 10^{-2}$  and  $\varepsilon = 10^{-2}$ ), the interval of admissible time steps  $[\Delta t_{min}, \Delta t_{max}]$  is larger than in the hyperbolic regime and as a result the *ImEx2-ctr* method does not generate any oscillations, as can be seen in Figures 7 and 8. We can also note that the *ImEx2-minmod* method yields equivalent performances to that of the *ImEx2-ctr* method, as expected from the theoretical study of the paper, while the classic *MUSCL-Hancock* method with splitting is already lagging behind in terms of accuracy.

**Acknowledgments.** We acknowledge the precious help of Loïc Gouarin and the use of the samurai code he develops (<https://github.com/hpc-maths/samurai>) within the framework of Initiative HPC@Maths (PI M. Massot and L. Gouarin - <https://initiative-hpc-maths.gitlab.labos.polytechnique.fr/site/>).

#### REFERENCES

- [1] G. ALBI, G. DIMARCO, AND L. PARESCHI, *Implicit-explicit multistep methods for hyperbolic systems with multiscale relaxation*, SIAM J. Sci. Comput., 42 (2020), pp. A2402–A2435.
- [2] A. ALVAREZ-LAGUNA, T. E. MAGIN, M. MASSOT, A. BOURDON, AND P. CHABERT, *Plasma-sheath transition in multi-fluid models with inertial terms under low pressure conditions: Comparison with the classical and kinetic theory*, Plasma Sources Science and Technology, 29 (2020), p. 025003.
- [3] A. ALVAREZ-LAGUNA, T. PICHARD, T. MAGIN, P. CHABERT, A. BOURDON, AND M. MASSOT, *An asymptotic preserving well-balanced scheme for the isothermal fluid equations in low-temperature plasma applications*, J. Comput. Phys., 419 (2020), p. 109634.
- [4] C. BARDOS, F. GOLSE, AND D. LEVERMORE, *Fluid dynamic limits of kinetic equations. I.*

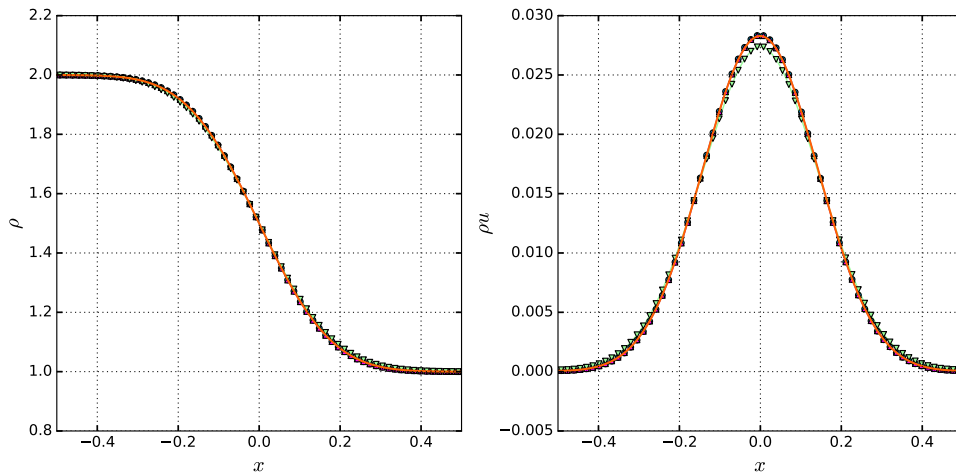


FIG. 8. Mass  $\rho$  and momentum  $\rho u$  obtained at time  $t_f = 0.01$  for  $\varepsilon = 10^{-2}$ ,  $\rho_L = 2$  and  $\rho_R = 1$ , respectively with the *ImEx2-ctr* method ( $\Delta$ ),  $\Delta t_{ctr} = 0.9\Delta t_{max}$ , from (5.11), *ImEx2-minmod* method ( $\circ$ ),  $\Delta t_{minmod} = 0.9\Delta t_{max}$ , from (5.11) and *MUSCL-Hancock* method with *Strang* splitting, *Reverse RK2* ( $\nabla$ ),  $\Delta t_{MH} = 0.9 \min(2\varepsilon^2/\sigma_{max}, \varepsilon\Delta x/(u_{max} + c))$ , using  $N = 64$  cells, with the latter scheme used as reference (solid line) with  $N = 2048$ .

- Formal derivations*, J. Stat. Phys., 63 (1991), pp. 323–344.
- [5] C. BERTHON AND R. TURPAULT, *Asymptotic preserving HLL schemes*, Numer. Methods Partial Differ. Equ., 27 (2011), pp. 1396–1422.
- [6] S. BOSCARINO, P. G. LEFLOCH, AND G. RUSSO, *High-order asymptotic-preserving methods for fully nonlinear relaxation problems*, SIAM J. Sci. Comput., 36 (2014), pp. A377–A395.
- [7] S. BOSCARINO, L. PARESCHI, AND G. RUSSO, *Implicit-explicit Runge-Kutta schemes for hyperbolic systems and kinetic equations in the diffusion limit*, SIAM J. Sci. Comput., 35 (2013), pp. A22–A51.
- [8] S. BOSCARINO AND G. RUSSO, *Flux-explicit IMEX Runge-Kutta schemes for hyperbolic to parabolic relaxation problems*, SIAM J. Numer. Anal., 51 (2013), pp. 163–190.
- [9] F. BOUCHUT, *Nonlinear stability of finite Volume Methods for hyperbolic conservation laws: And Well-Balanced schemes for sources*, Springer Science & Business Media, 2004.
- [10] C. BUET, B. DESPRÉS, AND E. FRANCK, *Design of asymptotic preserving finite volume schemes for the hyperbolic heat equation on unstructured meshes*, Numer. Math., 122 (2012), pp. 227–278.
- [11] J. C. BUTCHER, *Numerical methods for ordinary differential equations*, John Wiley & Sons, Ltd., Chichester, third ed., 2016. With a foreword by J. M. Sanz-Serna.
- [12] J. C. BUTCHER, *Numerical methods for ordinary differential equations*, John Wiley & Sons, Ltd., Chichester, third ed., 2016. With a foreword by J. M. Sanz-Serna.
- [13] R. E. CAFLISCH, S. JIN, AND G. RUSSO, *Uniformly accurate schemes for hyperbolic systems with relaxation*, SIAM J. Numer. Anal., 34 (1997), pp. 246–281.
- [14] J. R. CASH AND A. SINGHAL, *Mono-implicit Runge-Kutta Formulae for the Numerical Integration of Stiff Differential Systems*, IMA Journal of Numerical Analysis, 2 (1982), pp. 211–227.
- [15] C. CATTANEO, *Sulla conduzione del calore*, Atti del Seminario Matematico e Fisico dell’Università di Modena, 3 (1948), pp. 3–21.
- [16] C. CATTANEO, *Sur une forme de l’équation de la chaleur éliminant le paradoxe d’une propagation instantanée*, C. R. Acad. Sci. Paris, 247 (1958), pp. 431–433.
- [17] C. CHALONS AND R. TURPAULT, *High-order asymptotic-preserving schemes for linear systems: Application to the Goldstein-Taylor equations*, Numer. Methods Partial Differ. Equ., 35 (2019), pp. 1538–1561.
- [18] G.-Q. CHEN, C. D. LEVERMORE, AND T.-P. LIU, *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Commun. Pure Appl. Math., 47 (1994), pp. 787–830.

- [19] E. GODLEWSKI AND P.-A. RAVIART, *Numerical approximation of hyperbolic systems of conservation laws*, vol. 118, Springer Science & Business Media, 2013.
- [20] L. GOSSE AND G. TOSCANI, *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations*, C. R. Acad. Sci. Paris, 334 (2002), pp. 337–342.
- [21] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving ordinary differential equations. I*, vol. 8 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 1993. Nonstiff problems.
- [22] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving ordinary differential equations. I*, vol. 8 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 1993. Nonstiff problems.
- [23] S. JIN, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review*, Rivista di Matematica della Università di Parma, 3 (2012), pp. 177–216.
- [24] S. JIN AND C. D. LEVERMORE, *Numerical schemes for hyperbolic conservation laws with stiff relaxation terms*, J. Comput. Phys., 126 (1996), pp. 449–467.
- [25] S. JIN, L. PARESCHI, AND G. TOSCANI, *Diffusive relaxation schemes for multiscale discrete-velocity kinetic equations*, SIAM J. Numer. Anal., 35 (1998), pp. 2405–2439.
- [26] S. JIN, L. PARESCHI, AND G. TOSCANI, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM J. Numer. Anal., 38 (2000), pp. 913–936.
- [27] N. N. KALITKIN AND I. P. POSHIVAYLO, *Computations with inverse Runge-Kutta schemes*, Mathematical Models and Computer Simulations, 6 (2014), pp. 272–285.
- [28] N. N. KALITKIN AND I. P. POSHIVAYLO, *Computations with inverse Runge-Kutta schemes*, Mathematical Models and Computer Simulations, 6 (2014), pp. 272–285.
- [29] A. KLAR, *An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit*, SIAM J. Numer. Anal., 35 (1998), pp. 1073–1094.
- [30] P. D. LAX AND R. D. RICHTMYER, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math., 9 (1956), p. 267–293.
- [31] P. D. LAX AND B. WENDROFF, *Systems of conservation laws*, Comm. Pure Appl. Math., 13 (1960), p. 217–237.
- [32] J. C. MAXWELL, *On the dynamical theory of gases*, Philos. Trans. A: Math. Phys. Eng. Sci., 157 (1867), pp. 49–88.
- [33] R. NATALINI, *Convergence to equilibrium for the relaxation approximations of conservation laws*, Commun. Pure Appl. Math., 49 (1996), pp. 795–823.
- [34] L. REBOUL, *Mathematical modeling and simulation of non-equilibrium plasmas for the prediction of the electric propulsion*, PhD thesis, EDMH, Institut Polytechnique de Paris, 2023.
- [35] L. REBOUL, A. A. LAGUNA, T. MAGIN, A. BOURDON, AND M. MASSOT, *Fluid modeling of low-temperature partially-magnetized plasmas in the presence of sheaths: numerical issues for the simulation of instabilities and comparison with kinetic simulations*, Physics of Plasmas, (2022). To be submitted (a preliminary version has appeared as a preprint of the von Karman PhD symposium in 2021).
- [36] L. M. SKVORTSOV, *On implicit Runge–Kutta methods obtained as a result of the inversion of explicit methods*, Mathematical Models and Computer Simulations, 9 (2017), pp. 498–510.
- [37] L. M. SKVORTSOV, *On implicit Runge–Kutta methods obtained as a result of the inversion of explicit methods*, Mathematical Models and Computer Simulations, 9 (2017), pp. 498–510.
- [38] E. F. TORO, *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*, Springer Science & Business Media, 2013.
- [39] W. M. G. VAN BOKHOVEN, *Efficient higher order implicit one-step methods for integration of stiff differential equations*, BIT Numerical Mathematics, 20 (1980), pp. 34–43.
- [40] P. VERNOTTE, *Les paradoxes de la théorie continue de l'équation de la chaleur*, C. R. Acad. Sci. Paris, 246 (1958), pp. 3154–3155. and 1948 - Volume 227 - pages 43 and 114.
- [41] Y. ZHANG AND B. TABARROK, *Modifications to the Lax–Wendroff scheme for hyperbolic systems with source terms*, Int. J. Numer. Meth. Eng., 44 (1999), pp. 27–40.