



HAL
open science

On the stability of discrete boundaries conditions for linear hyperbolic systems

Geoffrey Beck, Nicolas Crouseilles, Ludovic Martaud

► **To cite this version:**

Geoffrey Beck, Nicolas Crouseilles, Ludovic Martaud. On the stability of discrete boundaries conditions for linear hyperbolic systems. 2025. ⟨hal-05008631v2⟩

HAL Id: hal-05008631

<https://hal.science/hal-05008631v2>

Preprint submitted on 6 Jan 2026

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.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

On the stability of discrete boundaries conditions for linear hyperbolic systems.

Geoffrey Beck*, Nicolas Crouseilles† Ludovic Martaud‡

December 19, 2025

Abstract

This work concerns the numerical approximations of the boundary conditions for linear hyperbolic systems. Discrete boundary procedure computations and an artificial viscosity are proposed to corrected any three-points finite volume schemes. The fully discrete stability of the resulting numerical schemes is established. Numerical test cases are performed to illustrate the relevancy of the proposed procedure.

Key words: Finite-volumes schemes, Boundary conditions, Discrete stability estimates, Kreiss symmetrizer.

1 Introduction

We investigate the stability of numerical schemes for linear initial hyperbolic boundary value problems cast on a half-line. Namely, denoting $|\cdot|_{\mathbb{R}^p}$ the standard euclidean norm of \mathbb{R}^p with $p \geq 1$ and $u(t, x) \in \mathbb{R}^p$ such that $\lim_{x \rightarrow +\infty} |u(t, x)|_{\mathbb{R}^p} = 0$, we are interested in the following Initial Boundary Value Problem (IBVP):

$$\begin{cases} \partial_t u + A \partial_x u = Qu, & \forall (t, x) \in (\mathbb{R}_*^+)^2, \\ Nu(t, 0) = g(t), & \forall t \geq 0, \\ u(t = 0, x) = u^{\text{in}}(x), & \forall x > 0, \end{cases} \quad (1.1)$$

where A is a constant \mathbb{R} -diagonalisable matrix in $\mathcal{M}_{p,p}(\mathbb{R})$, $Q \in \mathcal{M}_{p,p}(\mathbb{R})$. The initial data u^{in} is assumed to be regular enough, *i.e.* belongs to L^2 . The matrix N and the smooth data functions $g(t)$ prescribe the boundary conditions and they must satisfy some conditions ensuring the well-posedness of (1.1). In particular, the data satisfies on the edge $\{x = t = 0\}$ the following compatibility condition: $Nu^{\text{in}}(x = 0) = g(t = 0)$. The standard condition on N is called Kreiss-Lopatinskii [29] (see also [8, 15, 24, 34, 35] and the references within). In particular, such a condition imposes the number of boundary conditions to be equal to the

*Univ. Rennes, IRMAR UMR 6625 & Centre Inria de l'Université de Rennes (MINGuS) & ENS Rennes, France.

Email address: geoffrey.a.beck@inria.fr.

†Univ. Rennes, IRMAR UMR 6625 & Centre Inria de l'Université de Rennes (MINGuS) & ENS Rennes, France.

Email address: nicolas.crouseilles@inria.fr.

‡Univ. Rennes, IRMAR UMR 6625 & Centre Inria de l'Université de Rennes (MINGuS) & ENS Rennes, France.

Email address: ludovic.martaud@inria.fr.

number of positive eigenvalues of A . Moreover, the Kreiss-Lopatinskii condition implies a global L^2 -weighted estimate on $u(t, x)$ that formally writes

$$\frac{d}{dt} \int_{\mathbb{R}^+} (u^\top S u)(t, x) dx \lesssim |g(t)|, \quad (1.2)$$

where \cdot^\top denotes the usual transposition of \mathbb{R}^p and $S \in \mathcal{M}_{p,p}(\mathbb{R})$ is a symmetric definite positive matrix. For sake of simplicity, this formal energy estimate is valid for A invertible and $Q = 0$ (see Proposition 2.5 for the general case).

From a numerical point of view, the design of numerical schemes for (1.1) satisfying a fully discrete version of (1.2) is known to be challenging [11, 16–18]. The case of scalar equation, *i.e.* $p = 1$, is well studied (see [10–12] and the references within). For system of equations, *i.e.* $p > 1$, the main difficulties raise from the boundary conditions $Nu(t, 0) = g(t)$ that do not provide a complete knowledge of $u(t, 0)$ whereas any standard numerical scheme requires this knowledge. According to [19, 28] (see also the references within), the missing boundary conditions are defined by an admissible set and their numerical computation is for instance done with an exact resolution of the IBVP [19]. Another strategy consists in introducing the boundary condition in the approximation of $Au(n\Delta t, 0 + \Delta x/2)$ (with $\Delta t, \Delta x > 0$ and $n \in \mathbb{N}$) as follows:

$$Au(n\Delta t, 0 + \Delta x/2) \approx (A + \tilde{N} + \lambda_b)u_1^n - \tilde{N}u_0^n,$$

where $u_0^n \approx u(n\Delta t, 0)$, $u_1^n \approx u(n\Delta t, \Delta x)$ and where $(\lambda_b, \tilde{N}) \in \mathbb{R} \times \mathcal{M}_{p,p}(\mathbb{R})$ are chosen according to $Nu(t, 0) = g(t)$ and in order to enforce some stability properties. This strategy is derived within a suitable formulation of Discontinuous Galerkin schemes in [1, 2, 13] and for finite volume schemes in [14, 26]. The numerical boundary conditions may also be computed with a boundary layer correction based on a parabolic regularization of (1.1) combined with a asymptotic analysis that studies the vanishing regularization limit [19, 22]. On the other hand, the lacked boundary conditions are reachable from the Riemann invariants that correspond to a reformulation of (1.1) in the diagonal basis of A . In this regard, the Riemann invariants related to the negative eigenvalues of A are often approximated with a downwind discretization [7] then coupled to a three-points first-order finite volume scheme under the form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (\mathfrak{F}(u_i^n, u_{i+1}^n) - \mathfrak{F}(u_{i-1}^n, u_i^n)) + \Delta t Q u_i^n, \quad \forall i \geq 1, \quad (1.3)$$

where $\Delta t > 0$, $\mathfrak{F} : (\mathbb{R}^p)^2 \rightarrow \mathbb{R}^p$ is a consistent continuous numerical flux [25, 37, 38, 44] (see [32, 43] for detailed overviews and examples) and where $(u_i^n)_{i \geq 0}$ denotes a consistent approximation of $(u(n\Delta t, i\Delta x))_{i \geq 0}$, $\Delta x > 0$ being the space mesh step size.

The goal of this work is to enforce discrete stability properties for schemes (1.3) when the lacked boundary conditions are approximated with the discrete downwind Riemann invariants. Namely, following the strategy introduced in [39–42] (see also [9, 23]), we derive an artificial viscous correction for (1.3) in a such way the corrected scheme satisfies a fully discrete version of (1.2).

The paper is organized as follows. In Section 2, we propose a reformulation for the relation $Nu(t, 0) = g(t)$ and an explicit symmetrization for (1.1) to derive L^2 -weighted estimates under the form (1.2). Then, in Section 3, our numerical viscous correction for (1.3) and our main fully discrete stability result is given (Theorem 3.1). We would like to emphasize the constants involved in the continuous and discrete estimates are the same. Section 4 concerns the proof of our main result. Section 5 is devoted to the numerical experiments that illustrate our numerical procedure. In particular, one studies two specifics examples where one has explicit solutions to

make comparisons: the usual wave equations and relaxed dispersive waves. To catch relaxed dispersive waves, it was necessary to introduce the matrix Q and to consider the presence of zero eigenvalues of A case.

2 L^2 -weighted estimate for linear initial boundary value problems

This section is dedicated to the derivation a L^2 -weighted estimate for (1.1). According to [8, 24, 29], it is well-known that hyperbolic IBVP are well-posed if and only if the matrices (A, \mathbf{N}) satisfy the Kreiss-Lopatinskii condition. In this work, we assume a more restrictive assumption:

Hypothesis 2.1. (i) (*Hyperbolicity.*) One assumes that

$$A = PDP^{-1},$$

where D is the diagonal matrix that contains the real eigenvalues λ_i of A organized in descending order. One assumes that the matrix A admits eigenvalues of both signs and one defines

$$p^+ = \text{Card}(\{\lambda \in \text{Sp}(A) \mid \lambda > 0\}), \quad \pi^+ = \begin{pmatrix} I_{p^+} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{p,p}(\mathbb{R}), \quad (2.1a)$$

$$p^0 = \text{Card}(\{\lambda \in \text{Sp}(A) \mid \lambda = 0\}), \quad \pi^0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_{p^0} & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{p,p}(\mathbb{R}), \quad (2.1b)$$

$$p^- = \text{Card}(\{\lambda \in \text{Sp}(A) \mid \lambda < 0\}), \quad \pi^- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{p^-} \end{pmatrix} \in \mathcal{M}_{p,p}(\mathbb{R}), \quad (2.1c)$$

where $I_\alpha \in \mathcal{M}_{\alpha,\alpha}(\mathbb{R})$ denotes the identity matrix of size α and $\text{Sp}(A)$ the spectrum of A .

(ii) (*Lopatinskii condition.*) One assumes that $\mathbf{N} \in \mathcal{M}_{p^+,p}(\mathbb{R})$ (p^+ lines and p column) satisfies

$$\text{rank}(\mathbf{N}P\pi^+) = p^+.$$

The Hyperbolicity assumption 2.1-(i) is a standard assumption [8, 15] and the existence of positive eigenvalues is necessary to (1.1) defines an IBVP. In addition, the diagonalizable character of A enables to introduce the Riemann invariants

$$\varphi(t, x) = P^{-1}u(t, x), \quad (2.2)$$

that satisfies

$$\partial_t \varphi + D \partial_x \varphi = P^{-1}QP\varphi.$$

Lemma 2.2. *Let assume Hypothesis 2.1 and let consider the notation (2.2) for the Riemann invariants. Let us denote*

$$\tilde{g}(t) = \begin{pmatrix} g(t) \\ 0 \end{pmatrix} \in \mathbb{R}^p, \quad (2.3)$$

and $\tilde{\mathbf{N}} = \begin{pmatrix} \mathbf{N} \\ 0 \end{pmatrix} \in \mathcal{M}_{p,p}(\mathbb{R})$. Then there exists $B \in \mathcal{M}_{p,p}(\mathbb{R})$ such that, for all time t , the boundary conditions $\mathbf{N}u(t, 0) = g(t)$ on $\{x = 0\}$ are equivalent to

$$u(t, 0) = PB(\tilde{g}(t) + (\pi^0 + \pi^-)\varphi(t, 0)). \quad (2.4)$$

Proof. Inserting the Riemann invariants $\varphi = P^{-1}u$ in $Nu(t, 0) = g(t)$ and thanks to the definition of the projectors $(\pi^\alpha)_{\{+,0,-\}}$ given by (2.1), we have

$$\tilde{g}(t) = \tilde{N}P\varphi(t, 0) = \tilde{N}P\pi^+\varphi(t, 0) + \tilde{N}P(\pi^0 + \pi^-)\varphi(t, 0).$$

Adding $(\pi^0 + \pi^-)\varphi(t, 0)$ on both sides of the above equation, we deduce

$$(\tilde{N}P\pi^+ + \pi^0 + \pi^-)\varphi(t, 0) = \tilde{g}(t) + (I - \tilde{N}P)(\pi^0 + \pi^-)\varphi(t, 0).$$

According to Hypothesis 2.1-(ii), $\tilde{N}P\pi^+ + \pi^0 + \pi^-$ is necessary invertible so the above formulation is equivalent to

$$\varphi(t, 0) = (\tilde{N}P\pi^+ + \pi^0 + \pi^-)^{-1}(\tilde{g}(t) + (I - \tilde{N}P)(\pi^0 + \pi^-)\varphi(t, 0)).$$

Using $u = P\varphi$ in this last equation, we deduce the expected result, that concludes the proof. \square

Since (2.4) is equivalent to $\pi^+\varphi(0, t) = B(\tilde{g}(t) + (\pi^0 + \pi^-)\varphi(t, 0))$, above Lemma 2.2 highlights the different status of the Riemann invariants $\varphi(t, 0)$. The Riemann invariants $\pi^+\varphi(t, 0)$ related to the non-negative eigenvalues of A are entirely defined by the data $\tilde{g}(t)$ and by the Riemann invariants $(\pi^0 + \pi^-)\varphi(t, 0)$ related to the negative eigenvalues of A . The quantities $\tilde{g}(t)$ and $\varphi(t, 0)$ are coupled with the B matrix whom the existence is ensured by Lemma 2.2. The proof of the above lemma provides a procedure for the derivation of B and we give in Section 5 practical examples for its computation. In the sequel, we assume the conditions of Lemma 2.2 are fulfilled. As a consequence, the hyperbolic IBVP (1.1) now reads

$$\begin{cases} \partial_t u + A\partial_x u = Qu, & \forall (t, x) \in (\mathbb{R}_+^*)^2, \\ u(t, 0) = PB(\tilde{g}(t) + (\pi^0 + \pi^-)\varphi(t, 0)), & \forall t \geq 0, \\ u(t = 0, x) = u^{\text{in}}(x), & \forall x > 0. \end{cases} \quad (2.5)$$

Equipped with this definition, we now give estimates for (2.5). For the clarity in the forthcoming developments, $\|\cdot\|$ stands for the matrix norm induced by $|\cdot|_{\mathbb{R}^p}$ and we introduce $K \in \mathcal{M}_{p,p}(\mathbb{R})$ as follows:

$$K = P^{-1}QPB. \quad (2.6)$$

Our first following lemma concerns an estimate on the trace of the Riemann invariants associated to null eigenvalues.

Lemma 2.3. *Let assume Hypothesis 2.1. The Riemann invariants $\pi^0\varphi = \pi^0P^{-1}u$ related to the null eigenvalue of A satisfy*

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |\pi^0\varphi(t, 0)|_{\mathbb{R}^p}^2 &\leq \frac{1}{2} (\|\pi^0K\pi^+\| + 2\|\pi^0K\pi^0\| + \|\pi^0K\pi^-\|) |\pi^0\varphi(t, 0)|_{\mathbb{R}^p}^2 \\ &\quad + \frac{1}{2} \|\pi^0K\pi^+\| |\tilde{g}(t)|_{\mathbb{R}^p}^2 + \frac{1}{2} \|\pi^0K\pi^-\| |\pi^-\varphi(t, 0)|_{\mathbb{R}^p}^2, \end{aligned} \quad (2.7)$$

where $K \in \mathcal{M}_{p,p}(\mathbb{R})$ is given by (2.6).

Proof. By definition, the Riemann invariants related to the null eigenvalues of A are such that $\partial_t \pi^0\varphi = \pi^0P^{-1}Qu$. As a consequence, using the boundary conditions (2.4) and (2.6), we have

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} |\pi^0\varphi(t, 0)|_{\mathbb{R}^p}^2 \\ &= ((\pi^0\varphi)^\top \partial_t \pi^0\varphi)(t, 0), \\ &= (\pi^0\varphi(t, 0))^\top \pi^0P^{-1}QPB(\tilde{g}(t) + (\pi^0 + \pi^-)\varphi(t, 0)), \\ &= (\pi^0\varphi(t, 0))^\top (\pi^0K\pi^+)\tilde{g}(t) + (\pi^0\varphi(t, 0))^\top (\pi^0K\pi^0)\pi^0\varphi(t, 0) + (\pi^0\varphi(t, 0))^\top (\pi^0K\pi^-)\pi^-\varphi(t, 0). \end{aligned}$$

Now, using the Young inequality $a^\top A b \leq \|A\| (a^\top a + b^\top b)/2$ that holds for all matrices A and for all vectors a, b of relevant sizes with respect to A , the above equality becomes

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\pi^0 \varphi(t, 0)\|_{\mathbb{R}^p}^2 \\ & \leq \frac{1}{2} \|\pi^0 K \pi^+\| \left(|\tilde{g}(t)|_{\mathbb{R}^p}^2 + \|\pi^0 \varphi(t, 0)\|_{\mathbb{R}^p}^2 \right) + \|\pi^0 K \pi^0\| \|\pi^0 \varphi(t, 0)\|_{\mathbb{R}^p}^2, \\ & + \frac{1}{2} \|\pi^0 K \pi^-\| \left(\|\pi^0 \varphi(t, 0)\|_{\mathbb{R}^p}^2 + \|\pi^-\varphi(t, 0)\|_{\mathbb{R}^p}^2 \right), \end{aligned}$$

that concludes the proof. \square

We now give an L^2 -weighted estimate for the solution u of the IBVP (2.5). In this regard, we introduce the following Sobolev space:

$$L_S^2((\mathbb{R}^+)^2, \mathbb{R}^p) = \left\{ u : (\mathbb{R}^+)^2 \mapsto \mathbb{R}^p \mid \int_{\mathbb{R}^+} (u^\top S u)(t, x) dx < +\infty \right\}, \quad (2.8)$$

where $S \in \mathcal{M}_{p,p}(\mathbb{R})$ is a symmetric definite positive matrix. In this work, we propose to choose S such that

$$S_m = P^{-\top} (\pi^+ + m(\pi^0 + \pi^-)) P^{-1}, \quad (2.9)$$

where $m > 0$ and where the diagonal projectors $(\pi^\alpha)_{\alpha \in \{+, 0, -\}}$ are defined in (2.1). The above matrix S_m Friedrich-symmetrizes (2.5), i.e. $(S_m A)^\top = S_m A$ and the goal is to constrain $m > 0$ in (2.9) to reach a Kreiss symmetrization of (2.5). Such a strategy is not new [27, 31], but we recall it because the discrete stability proposed in this work will mimic the same approach. The first following lemma establishes an estimate for the energy $\mathcal{E}(t)$ defined as

$$\mathcal{E}(t) := \int_{\mathbb{R}^+} (u^\top S_m u)(t, x) dx. \quad (2.10)$$

Lemma 2.4. *Let assume Hypothesis 2.1. Let $m > 0$ and $S_m \in \mathcal{M}_{p,p}(\mathbb{R})$ reads under the form (2.9). For all $m > 0$ and for all $u \in L_{S_m}^2((\mathbb{R}^+)^2, \mathbb{R}^p)$ solution of (2.5), we have the following statements*

(i) *The energy $\mathcal{E}(t)$ defined in (2.10) also writes*

$$\mathcal{E}(t) = \int_{\mathbb{R}^+} |\pi^+ \varphi(t, x)|_{\mathbb{R}^p}^2 dx + m \int_{\mathbb{R}^+} |(\pi^0 + \pi^-) \varphi(t, x)|_{\mathbb{R}^p}^2 dx, \quad (2.11)$$

where $\varphi = P^{-1}u$ denote the Riemann invariants and it satisfies

(ii)

$$\frac{1}{2} \frac{d}{dt} \mathcal{E}(t) \leq \frac{1}{2} (u^\top S_m A u)(t, 0) + \gamma_m \mathcal{E}(t), \quad (2.12)$$

where $\gamma_m \geq 0$ is given by

$$\gamma_m = \left(1 + \frac{1}{2\sqrt{m}} + \frac{\sqrt{m}}{2} \right) \|P^{-1} Q P\|. \quad (2.13)$$

Proof. Since (i) is directly obtained by inserting (2.9) in (2.10), we focus on the proof of (ii). Multiplying (2.5) by $u^\top S_m$ and integrating on $x \in [0, +\infty)$ leads to

$$\int_{\mathbb{R}^+} (u^\top S_m (\partial_t u + A \partial_x u))(t, x) dx = \int_{\mathbb{R}^+} (u^\top S_m Q u)(t, x) dx.$$

Using the definition of $\mathcal{E}(t)$ given by (2.10) and since $S_m^\top = S_m$, $(S_m A)^\top = S_m A$, the above formulation writes

$$\frac{1}{2} \frac{d}{dt} \mathcal{E}(t) = \frac{1}{2} (u^\top S_m A u)(t, 0) + \int_{\mathbb{R}^+} (u^\top S_m Q u)(t, x) dx. \quad (2.14)$$

Now, let consider the last term. From the definition (2.9) of S_m and since the Riemann invariants write $P^{-1}u = \varphi = \pi^+ \varphi + (\pi^0 + \pi^-) \varphi$, we get

$$\begin{aligned} u^\top S_m Q u &= \varphi^\top (\pi^+ + m(\pi^0 + \pi^-)) P^{-1} Q P (\pi^+ \varphi + (\pi^0 + \pi^-) \varphi), \\ &= (\pi^+ \varphi)^\top P^{-1} Q P (\pi^+ \varphi) + \frac{1}{\sqrt{m}} (\pi^+ \varphi)^\top P^{-1} Q P (\sqrt{m}(\pi^0 + \pi^-) \varphi) \\ &\quad + \sqrt{m} (\sqrt{m}(\pi^0 + \pi^-) \varphi)^\top P^{-1} Q P (\pi^+ \varphi) + m ((\pi^0 + \pi^-) \varphi)^\top P^{-1} Q P ((\pi^0 + \pi^-) \varphi). \end{aligned} \quad (2.15)$$

Arguing the Young inequality and observing $|\pi^+ \varphi|_{\mathbb{R}^p}^2 + m |(\pi^0 + \pi^-) \varphi|_{\mathbb{R}^p}^2 = u^\top S_m u$, we deduce from the above identity:

$$\begin{aligned} u^\top S_m Q u &\leq \|P^{-1} Q P\| |\pi^+ \varphi|_{\mathbb{R}^p}^2 + \frac{1}{2\sqrt{m}} \|P^{-1} Q P\| (|\pi^+ \varphi|_{\mathbb{R}^p}^2 + m |(\pi^0 + \pi^-) \varphi|_{\mathbb{R}^p}^2) \\ &\quad + \frac{\sqrt{m}}{2} \|P^{-1} Q P\| (|\pi^+ \varphi|_{\mathbb{R}^p}^2 + m |(\pi^0 + \pi^-) \varphi|_{\mathbb{R}^p}^2) + m \|P^{-1} Q P\| |(\pi^0 + \pi^-) \varphi|_{\mathbb{R}^p}^2, \\ &\leq \left(1 + \frac{1}{2\sqrt{m}} + \frac{\sqrt{m}}{2}\right) \|P^{-1} Q P\| u^\top S_m u. \end{aligned} \quad (2.16)$$

As a consequence, using the above upper-bound and the definition of $\gamma_m \geq 0$ given by (2.13), (2.14) becomes

$$\frac{1}{2} \frac{d}{dt} \mathcal{E}(t) \leq \frac{1}{2} (u^\top S_m A u)(t, 0) + \gamma_m \mathcal{E}(t). \quad (2.17)$$

□

It is worth noticing that if $u^\top S_m Q u \leq 0$, then the constant γ_m in (2.12) vanishes. The above lemma and Lemma 2.3 enable to derive the L^2 -weighted estimate that is given under the form defined by Kreiss in [29] (see also [8, 15, 24, 34, 35] and the references within). This estimate is reached by constraints on $m > 0$ in (2.9).

Proposition 2.5 (Kreiss type inequality for (2.5)). *Let $\alpha, C_\alpha, C'_\alpha > 0$ and $\text{Kr}_m \in \mathbb{R}$ such that*

$$C_\alpha = \max_{i \in [1, p]} \max(0, \lambda_i) (1 + \alpha) \|\pi^+ B \pi^+\|^2 + \frac{1}{2} \|\pi^0 K \pi^+\|, \quad (2.18a)$$

$$\begin{aligned} C'_\alpha &= \max_{i \in [1, p]} \max(0, \lambda_i) \left(1 + \frac{1}{\alpha}\right) \|\pi^+ B \pi^0\|^2 + \frac{1}{2} \|\pi^0 K \pi^+\| + \|\pi^0 K \pi^0\| \\ &\quad + \frac{1}{2} \|\pi^0 K \pi^-\|, \end{aligned} \quad (2.18b)$$

$$-\text{Kr}_m = \max_{i \in [1, p]} \max(0, \lambda_i) \|\pi^+ B \pi^-\|^2 + \frac{1}{2} \|\pi^0 K \pi^-\| + \frac{m}{2} \max_{i \in [1, p]} \min(0, \lambda_i). \quad (2.18c)$$

If $m \geq 0$ is such that $\text{Kr}_m \geq 0$, then the following holds:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} (|\pi^0 \varphi(t, 0)|_{\mathbb{R}^p}^2 + \mathcal{E}(t)) &+ \text{Kr}_m |\pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2 \\ &\leq C_\alpha |\tilde{g}(t)|_{\mathbb{R}^p}^2 + C'_\alpha |\pi^0 \varphi(t, 0)|_{\mathbb{R}^p}^2 + \gamma_m \mathcal{E}(t), \end{aligned} \quad (2.19)$$

where $\mathcal{E}(t)$, $\gamma_m \geq 0$ are defined in (2.10), (2.13) respectively.

The constants $C_\alpha, C'_\alpha > 0$ defined in (2.18) depend on $\alpha > 0$ that is an additional parameter. This parameter is left free for the trade off minimization of constants C_α and C'_α . For instance, if A, N are such that $\pi^+ B \pi^0 = 0$ then $\alpha = 0$ is an admissible choice. The prove of the above proposition now requires a Kreiss-like symmetrizer for (2.5). Such a symmetrization is detailed below.

Lemma 2.6 (Kreiss-like symmetrizer for (2.5)). *Let $\alpha, C_\alpha, C'_\alpha > 0$ and $\text{Kr}_m \in \mathbb{R}$ given in proposition 2.5. One has*

$$\begin{aligned} (u^\top S_m A u)(t, 0) &\leq - \left(\text{Kr}_m - \frac{1}{2} \|\|\| \pi^0 K \pi^- \|\|\| \right) |\pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2 + \left(C_\alpha - \frac{1}{2} \|\|\| \pi^0 K \pi^+ \|\|\| \right) |\tilde{g}(t)|_{\mathbb{R}^p}^2 \\ &\quad + \left(C'_\alpha - \frac{1}{2} \|\|\| \pi^0 K \pi^+ \|\|\| - \|\|\| \pi^0 K \pi^0 \|\|\| \right) |\pi^0 \varphi(t, 0)|_{\mathbb{R}^p}^2. \end{aligned} \quad (2.20)$$

Proof. Using the spectral decomposition $A = P D P^{-1}$, the definition (2.9) of S_m and $\varphi = P^{-1} u$, we have

$$\begin{aligned} (u^\top S_m A u)(t, 0) &= (u^\top P^{-\top} (\pi^+ + m(\pi^0 + \pi^-)) P^{-1} P D P^{-1} u)(t, 0), \\ &= (\varphi^\top \pi^+ D \varphi)(t, 0) + m(\varphi^\top \pi^0 D \varphi)(t, 0) + m(\varphi^\top \pi^- D \varphi)(t, 0), \\ &= (\varphi^\top \pi^+ D \varphi)(t, 0) + m(\varphi^\top \pi^- D \varphi)(t, 0). \end{aligned}$$

Since D contains the eigenvalues of A organized in descending order, we have

$$\begin{aligned} &(\varphi^\top \pi^+ D \varphi)(t, 0) + m(\varphi^\top \pi^- D \varphi)(t, 0) \\ &\leq \max_{i \in [1, p]} \max(0, \lambda_i) |\pi^+ \varphi(t, 0)|_{\mathbb{R}^p}^2 + m \left(\max_{i \in [1, p]} \min(0, \lambda_i) \right) |\pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2. \end{aligned} \quad (2.21)$$

But, using (2.4) to reformulate $\pi^+ \varphi(t, 0)$, then arguing the triangle inequality and the Young inequality $2ab \leq \alpha a^2 + b^2/\alpha$ that holds for all $(a, b, \alpha) \in \mathbb{R}^2 \times \mathbb{R}_*^+$, we deduce

$$\begin{aligned} |\pi^+ \varphi(t, 0)|_{\mathbb{R}^p}^2 &= |\pi^+ B (\tilde{g}(t) + (\pi^0 + \pi^-) \varphi(t, 0))|_{\mathbb{R}^p}^2, \\ &\leq 2 |\pi^+ B (\tilde{g}(t) + \pi^0 \varphi(t, 0))|_{\mathbb{R}^p}^2 + 2 |\pi^+ B \pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2, \\ &\leq 2(1 + \alpha) |\pi^+ B \tilde{g}(t)|_{\mathbb{R}^p}^2 + 2 \left(1 + \frac{1}{\alpha}\right) |\pi^+ B \pi^0 \varphi(t, 0)|_{\mathbb{R}^p}^2 + 2 \|\|\| \pi^+ B \pi^- \|\|\|^2 |\pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2, \\ &\leq 2(1 + \alpha) \|\|\| \pi^+ B \pi^+ \|\|\|^2 |\tilde{g}(t)|_{\mathbb{R}^p}^2 + 2 \left(1 + \frac{1}{\alpha}\right) \|\|\| \pi^+ B \pi^0 \|\|\|^2 |\pi^0 \varphi(t, 0)|_{\mathbb{R}^p}^2 \\ &\quad + 2 \|\|\| \pi^+ B \pi^- \|\|\|^2 |\pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2. \end{aligned} \quad (2.22)$$

As a consequence, inserting (2.22) in (2.21), one has (2.20). \square

Equipped with the above Kreiss-like symmetrizer for (2.5), the proof of Proposition 2.5 is now given.

Proof of Proposition 2.5. Inserting (2.20) in (2.12), one has the following estimation of the energy

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \mathcal{E}(t) &\leq \max_{i \in [1, p]} \max(0, \lambda_i) \left((1 + \alpha) \|\|\| \pi^+ B \pi^+ \|\|\|^2 |\tilde{g}(t)|_{\mathbb{R}^p}^2 + \left(1 + \frac{1}{\alpha}\right) \|\|\| \pi^+ B \pi^0 \|\|\|^2 |\pi^0 \varphi(t, 0)|_{\mathbb{R}^p}^2 \right) \\ &\quad + \left(\max_{i \in [1, p]} \max(0, \lambda_i) \|\|\| \pi^+ B \pi^- \|\|\|^2 + \frac{m}{2} \max_{i \in [1, p]} \min(0, \lambda_i) \right) |\pi^- \varphi(t, 0)|_{\mathbb{R}^p}^2 + \gamma_m \mathcal{E}(t), \end{aligned}$$

The initial sequence $(u_i^0)_{i \geq 1}$ is given by $u_i^0 = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u^{\text{in}}(x) dx$, for all $i \geq 1$. From the sequence $(u_i^n)_{i \geq 1}$, we propose to compute the updated sequence $(u_i^{n+1})_{i \geq 1}$ with a viscous version of a generic three-points explicit numerical scheme (1.3). Namely, we consider finite volume scheme under the form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} \left(\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n) \right) + \Delta t Q u_i^n, \quad \forall i \geq 1, \quad (3.2)$$

where $\mathcal{F} : (\mathbb{R}^p)^2 \rightarrow \mathbb{R}^p$ is given by

$$\mathcal{F}(u_i^n, u_{i+1}^n) = \mathfrak{F}(u_i^n, u_{i+1}^n) - \frac{1}{2} P (\nu_{i+\frac{1}{2}}^n + |D|) P^{-1} (u_{i+1}^n - u_i^n), \quad \forall i \geq 0. \quad (3.3)$$

In the above definition, the second term stands for the artificial viscosity where $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0} \in \mathcal{M}_{p,p}(\mathbb{R})$ is a sequence of diagonal matrices, $|D|$ denotes the diagonal matrices that contains the absolute value of the eigenvalues of A organized in descending order. According to [9, 23], the artificial viscosity term can be added to any consistent numerical flux $\mathfrak{F} : (\mathbb{R}^p)^2 \rightarrow \mathbb{R}^p$ that satisfies the consistency relation $\mathfrak{F}(u, u) = Au$. The main originality of this work is to consider a non-homogeneous and anisotropic viscosity to enforce stability properties in the presence of boundaries conditions $u(t, 0)$ defined in Lemma 2.2. The discrete version of these boundaries conditions is denoted $u_0^n \in \mathbb{R}^p$ and it now requires a suitable definition to complete the viscous scheme (3.2).

According to Lemma 2.2 (see also Figure 1) u_0^n is such that

$$u_0^n = PB(\tilde{g}^n + (\pi^0 + \pi^-)\varphi_0^n), \quad \forall n \geq 0. \quad (3.4)$$

Therefore, u_0^n is defined by the knowledge of $(\tilde{g}^n, (\pi^0 + \pi^-)\varphi_0^n) \in (\mathbb{R}^p)^2$ that stands for a consistent discretization of the boundary data $\tilde{g}(t^n)$ and of the Riemann invariants $(\pi^0 + \pi^-)\varphi(t^n, 0)$ satisfying

$$(\pi^0 + \pi^-)(\partial_t \varphi)(t^n, 0) + (\pi^0 + \pi^-)D(\partial_x \varphi)(t^n, 0) = (\pi^0 + \pi^-)P^{-1}Qu(t^n, 0). \quad (3.5)$$

In this work, $\tilde{g}^n \in \mathbb{R}^p$ is left free, but we define the discrete Riemann invariants $(\pi^0 + \pi^-)\varphi_0^n$ as follows:

$$\begin{aligned} (\pi^0 + \pi^-)\varphi_0^0 &= (\pi^0 + \pi^-)P^{-1}u_1^0, \\ (\pi^0 + \pi^-)\varphi_0^{n+1} &= (\pi^0 + \pi^-)\varphi_0^n - \frac{\Delta t}{\Delta x} (\pi^0 + \pi^-)D(\varphi_1^n - \varphi_0^n) \\ &\quad + \Delta t (\pi^0 + \pi^-)P^{-1}Qu_0^n, \quad \forall n \geq 0, \end{aligned} \quad (3.6)$$

The formulation (3.6) is consistent with (3.5). If the data are regular, we also have

$$(\pi^0 + \pi^-)P^{-1}u_1^0 = (\pi^0 + \pi^-)P^{-1}u_0^0.$$

Our main result given below concerns a fully discrete L^2 -estimate for the scheme (3.2), (3.4), (3.6).

Theorem 3.1 (Fully discrete L^2 -weighted estimate for (3.2), (3.4), (3.6)). *Let us consider the IBVP (2.5) endowed by a symmetrizer $S_m \in \mathcal{M}_{p,p}(\mathbb{R})$ of the form (2.9) with $m > 0$ such that $\text{Kr}_m \geq 0$ in (2.18c). Furthermore, let assume Hypothesis 2.1. Let $(u_i^n)_{i \geq 0}$ a non-constant sequence such that $\sum_{i \geq 0} (u_i^n)^\top S_m u_i^n \Delta x < +\infty$ and let*

$$\mathcal{E}^n = \sum_{i \geq 1} |\pi^+ \varphi_i^n|_{\mathbb{R}^p}^2 \Delta x + m \sum_{i \geq 0} |(\pi^0 + \pi^-)\varphi_i^n|_{\mathbb{R}^p}^2 \Delta x, \quad \text{with } \varphi_i^n = Pu_i^n. \quad (3.7)$$

Let us also consider a viscous scheme (3.2) and assume the existence of bounded diagonal matrices $k : (\mathbb{R}^p)^2 \rightarrow \mathcal{M}_{p,p}(\mathbb{R})$ satisfying

$$2(v - u)^\top S_m \left(\mathfrak{F}(u, v) - \frac{1}{2}A(u + v) \right) \leq (v - u)^\top S_m P k(u, v) P^{-1} (v - u), \quad (3.8)$$

for all $(u, v) \in (\mathbb{R}^p)^2$. If $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ in (3.3) is such that

$$\begin{aligned} & (u_1^n)^\top S_m \left(\mathfrak{F}(u_0^n, u_1^n) - \frac{1}{2}A(u_1^n + u_0^n) - \frac{1}{2}P\nu_{\frac{1}{2}}^n P^{-1}(u_1^n - u_0^n) \right) \\ & + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P \left(-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n) \right) P^{-1} (u_{i+1}^n - u_i^n) \\ & - \frac{1}{2} \sum_{i \geq 0} (u_{i+1}^n - u_i^n)^\top S_m P |D| P^{-1} (u_{i+1}^n - u_i^n) < 0, \end{aligned} \quad (3.9)$$

then there exists (explicit) $\frac{\Delta t}{\Delta x} > 0$ small enough such that $(\mathcal{E}^{n+1}, (\pi^0 + \pi^-)\varphi_0^{n+1})$ given by (3.2), (3.6) satisfy:

$$\begin{aligned} & \frac{|\pi^0 \varphi_0^{n+1}|_{\mathbb{R}^p}^2 - |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2}{2\Delta t} + \frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{2\Delta t} + \text{Kr}_m |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2 \\ & \leq \left(C_\alpha + \frac{m\Delta x}{2} \|\|(\pi^0 + \pi^-)K\pi^+\|\| \right) |\tilde{g}^n|_{\mathbb{R}^p}^2 + C'_\alpha |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + C''_m \mathcal{E}^n, \end{aligned} \quad (3.10)$$

with $K = P^{-1}QP$ given by (2.6), $C_\alpha, C'_\alpha > 0$ and $\text{Kr}_m, \gamma_m \geq 0$ are given by (2.18), (2.13) and where

$$C''_m = \max \left(\gamma_m, \frac{1}{2} \|\|(\pi^0 + \pi^-)K\pi^+\|\| + \|\|(\pi^0 + \pi^-)K(\pi^0 + \pi^-)\|\| \right). \quad (3.11)$$

We recall that if $u^\top S_m Q u \leq 0$, then the constant γ_m in (3.11) vanishes. Up to the term $m\Delta x \|\|(\pi^0 + \pi^-)K\pi^+\|\|/2 \geq 0$, (3.10) is exactly a discrete version of (2.19). Since $\text{Kr}_m \geq 0$ and $m\Delta x \|\|(\pi^0 + \pi^-)K\pi^+\|\|/2$ is bounded as $\Delta x \rightarrow 0$, (3.10) is in a suitable form to apply a discrete Grönwall's inequality to the sequence $(|\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + \mathcal{E}^n)_{n \geq 0}$. It also ensures an estimate for $(u_i^n)_{i \geq 1}$ because, according to the definitions of $(S_m, \mathcal{E}^n) \in \mathcal{M}_{p,p}(\mathbb{R}) \times \mathbb{R}^+$, given by (2.9), (3.7), we have

$$\begin{aligned} \sum_{i \geq 1} (u_i^n)^\top S_m u_i^n \Delta x &= \sum_{i \geq 1} \left(|\pi^+ \varphi_i^n|_{\mathbb{R}^p}^2 + m |(\pi^0 + \pi^-) \varphi_i^n|_{\mathbb{R}^p}^2 \right) \Delta x, \\ &\leq m\Delta x |\pi^0 + \pi^-|_{\mathbb{R}^p}^2 + \sum_{i \geq 1} \left(|\pi^+ \varphi_i^n|_{\mathbb{R}^p}^2 + m |(\pi^0 + \pi^-) \varphi_i^n|_{\mathbb{R}^p}^2 \right) \Delta x, \\ &\leq \sum_{i \geq 1} |\pi^+ \varphi_i^n|_{\mathbb{R}^p}^2 \Delta x + m \sum_{i \geq 0} |(\pi^0 + \pi^-) \varphi_i^n|_{\mathbb{R}^p}^2 \Delta x, \\ &\leq \mathcal{E}^n. \end{aligned} \quad (3.12)$$

Hence, we deduce the existence of $C_n > 0$ bounded as $\Delta x, \Delta t \rightarrow 0$ and such that

$$|\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + \mathcal{E}^n \leq C_n \left(|\pi^0 \varphi_0^0|_{\mathbb{R}^p}^2 + \mathcal{E}^0 + \sum_{k=0}^{n-1} |\tilde{g}^k|_{\mathbb{R}^p}^2 \Delta t \right), \quad \forall n \geq 1.$$

It is also worth noting that the summations in (3.7) do not start at the same index ($i \geq 1$ for the first sum and $i \geq 0$ for the second sum). Indeed, \mathcal{E}^n defined in (3.7) is consistent with (2.11) but the two different starting indices in (3.7) highlight the different status of the discrete Riemann

invariants : $(\pi^+ \varphi_i^{n+1})_{i \geq 1}$ is only computed by (3.2) whereas $((\pi^0 + \pi^-) \varphi_i^{n+1})_{i \geq 0}$ is computed by (3.2) and (3.6). These different status may cause numerical instabilities (see Section 5) and thus motivate the viscous correction in (3.3).

This viscous correction is given in Theorem 3.1 under a Lipschitz continuity (3.8) for $\mathfrak{F}(u, v)$. This assumption is satisfied by several first-order numerical flux [25, 37, 38, 44] and the reader may refer to [32, 43] for detailed reviews. As a consequence, Theorem 3.1 shows that if the artificial viscous correction $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ is chosen according to (3.9), the discrete boundary (3.4) may be used to complete any numerical flux $\mathfrak{F}(u_i^n, u_{i+1}^n)$ under the form (3.3). The inequality (3.9) defines a constraint to choose the sequence $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$. For a non constant sequence $(u_i^n)_{i \geq 0}$, since the last term in (3.9) is negative, it is actually sufficient to ensure

$$(u_1^n)^\top S_m \left(\mathfrak{F}(u_0^n, u_1^n) - \frac{1}{2} A(u_1^n + u_0^n) - P \nu_{\frac{1}{2}}^n P^{-1} (u_1^n - u_0^n) \right) \leq 0, \quad (3.13)$$

and

$$\frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P \left(-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n) \right) P^{-1} (u_{i+1}^n - u_i^n) \leq 0. \quad (3.14)$$

Enforcing (3.13) and (3.14) thanks to a suitable choice of $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ and for $\Delta t > 0$ is discussed in Section 5.

The scheme is consistant and stable. However, one cannot directly apply the Lax equivalence Theorem to show that the scheme is convergent because it is not necessary linear. For example, in the numerical simulations given in Section 5, the procedure for the boundary given by (5.3) is non-linear. Now, the goal is to prove Theorem 3.1.

4 Discrete stability analysis

This section concerns the proof of Theorem 3.1 which is organized as follows. In Subsection 4.1, we derive estimates for the discrete Riemann invariants $\pi^0 \varphi_0^{n+1}$ and $(\pi^0 + \pi^-) \varphi_0^{n+1}$ defined in (3.6). We secondly use (3.2) to derive discrete estimates for $(u_i^{n+1})_{i \geq 1}$ in Subsection 4.2. Subsection 4.3 uses Subsections 4.1 and 4.2 to define a suitable upper-bound for the norm of $((\pi^0 + \pi^-) \varphi_0^{n+1}, (u_i^{n+1})_{i \geq 1})$. Then, in Subsection 4.4, the proof of Theorem 3.1 is given.

4.1 Estimates for the discrete Riemann invariants in the boundary cell

This section concerns the derivation of estimates for the discrete Riemann invariants $\pi^0 \varphi_0^n$ and $(\pi^0 + \pi^-) \varphi_0^n$, both defined in (3.6). Since $\pi^0 \pi^- = 0$ and $\pi^0 D = 0$, (3.6) leads to the following formulation for $\pi^0 \varphi_0^n$:

$$\frac{\pi^0 \varphi_0^{n+1} - \pi^0 \varphi_0^n}{\Delta t} = \pi^0 P^{-1} Q u_0^n, \quad (4.1)$$

where $u_0^n \in \mathbb{R}^p$ is defined in (3.4). The first following lemma is a discrete version of Lemma 2.3 and it derives a quadratic estimate for the above updated quantity $\pi^0 \varphi_0^{n+1}$.

Lemma 4.1. *Let us consider the discrete Riemann invariants $\pi^0 \varphi_0^n$ defined by (4.1). The updated discrete Riemann invariant $\pi^0 \varphi_0^{n+1}$ satisfies*

$$\begin{aligned} & \frac{|\pi^0 \varphi_0^{n+1}|_{\mathbb{R}^p}^2 - |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2}{2\Delta t} \\ & \leq \frac{1}{2} \left(\|\pi^0 K \pi^+\| + 2\|\pi^0 K \pi^0\| + \|\pi^0 K \pi^-\| \right) |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 \\ & \quad + \frac{1}{2} \|\pi^0 K \pi^+\| |\tilde{g}^n|_{\mathbb{R}^p}^2 + \frac{1}{2} \|\pi^0 K \pi^-\| |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2 + \frac{\Delta t}{2} |\pi^0 P^{-1} Q u_0^n|_{\mathbb{R}^p}^2. \end{aligned} \quad (4.2)$$

where $K \in \mathcal{M}_{p,p}(\mathbb{R})$ is given by (2.6).

Proof. According to (4.1), and using the formulation of u_0^n given by (3.4), we have

$$|\pi^0 \varphi_0^{n+1}|_{\mathbb{R}^p}^2 = |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + 2\Delta t (\varphi_0^n)^\top \pi^0 K (\tilde{g}^n + (\pi^0 + \pi^-) \varphi_0^n) + \Delta t^2 |\pi^0 P^{-1} Q u_0^n|_{\mathbb{R}^p}^2. \quad (4.3)$$

Recalling the definition (2.3) of \tilde{g} (which implies $\pi^+ \tilde{g} = \tilde{g}$) and using the Young inequality, we have for the second term of the right hand-side

$$\begin{aligned} & (\varphi_0^n)^\top \pi^0 K (\tilde{g}^n + (\pi^0 + \pi^-) \varphi_0^n) \\ &= (\pi^0 \varphi_0^n)^\top (\pi^0 K \pi^+) \tilde{g}^n + (\pi^0 \varphi_0^n)^\top (\pi^0 K \pi^0) \pi^0 \varphi_0^n + (\pi^0 \varphi_0^n)^\top (\pi^0 K \pi^-) \pi^- \varphi_0^n, \\ &\leq \frac{1}{2} \|\pi^0 K \pi^+\| (|\tilde{g}^n|_{\mathbb{R}^p}^2 + |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2) + \|\pi^0 K \pi^0\| |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + \frac{1}{2} \|\pi^0 K \pi^-\| (|\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2). \end{aligned}$$

Associating the above formulation to (4.3) enables to get the expected result. \square

The second following lemma concerns an estimate for the updated quantity $(\pi^0 + \pi^-) \varphi_0^{n+1}$ defined by (3.6).

Lemma 4.2. *Let us consider the discrete Riemann invariants $(\pi^0 + \pi^-) \varphi_0^n$ defined by (3.6). For all $m > 0, \Delta x > 0$, the updated discrete Riemann invariants $(\pi^0 + \pi^-) \varphi_0^{n+1}$ satisfy*

$$\begin{aligned} & m\Delta x \frac{|\pi^0 + \pi^- \varphi_0^{n+1}|_{\mathbb{R}^p}^2 - |(\pi^0 + \pi^-) \varphi_0^n|_{\mathbb{R}^p}^2}{2\Delta t} \\ &\leq -m(\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) \\ &\quad + C_m'' m |(\pi^0 + \pi^-) \varphi_0^n|_{\mathbb{R}^p}^2 \Delta x + \frac{m\Delta x}{2} \|\pi^0 + \pi^-\| K \pi^+ \|\tilde{g}^n\|_{\mathbb{R}^p}^2 \\ &\quad + \frac{m\Delta t \Delta x}{2} |(\pi^0 + \pi^-) \left(-D \frac{\varphi_1^n - \varphi_0^n}{\Delta x} + P^{-1} Q u_0^n \right)|_{\mathbb{R}^p}^2, \end{aligned} \quad (4.4)$$

where $C_m'' \geq 0, K \in \mathcal{M}_{p,p}(\mathbb{R})$ are given by (3.11), (2.6) respectively.

Proof. Applying $((\pi^0 + \pi^-) \varphi_0^n)^\top$ to (3.6), since $(\pi^0 + \pi^-)^2 = (\pi^0 + \pi^-)$ and using the identity $v \cdot u - |v|^2 = |u|^2/2 - |v|^2/2 - |v - u|^2/2$ with $u = \varphi_0^{n+1}, v = \varphi_0^n$, we get

$$\begin{aligned} & |(\pi^0 + \pi^-) \varphi_0^{n+1}|_{\mathbb{R}^p}^2 \\ &= |(\pi^0 + \pi^-) \varphi_0^n|_{\mathbb{R}^p}^2 - 2 \frac{\Delta t}{\Delta x} (\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) + 2\Delta t (\varphi_0^n)^\top (\pi^0 + \pi^-) P^{-1} Q u_0^n \\ &\quad + \Delta t^2 |(\pi^0 + \pi^-) \left(-D \frac{\varphi_1^n - \varphi_0^n}{\Delta x} + P^{-1} Q u_0^n \right)|_{\mathbb{R}^p}^2. \end{aligned}$$

Reorganizing the terms and multiplying both sides by $\frac{m\Delta x}{2\Delta t} > 0$ leads to

$$\begin{aligned} & m\Delta x \frac{|\pi^0 + \pi^- \varphi_0^{n+1}|_{\mathbb{R}^p}^2 - |(\pi^0 + \pi^-) \varphi_0^n|_{\mathbb{R}^p}^2}{2\Delta t} \\ &= -m(\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) + m\Delta x (\varphi_0^n)^\top (\pi^0 + \pi^-) P^{-1} Q u_0^n \\ &\quad + \frac{m\Delta x \Delta t}{2} |(\pi^0 + \pi^-) \left(-D \frac{\varphi_1^n - \varphi_0^n}{\Delta x} + P^{-1} Q u_0^n \right)|_{\mathbb{R}^p}^2. \end{aligned} \quad (4.5)$$

Let now consider the second term on the right hand-side. Using (3.4) to reformulate u_0^n and recalling $K = P^{-1} Q P B$, we get

$$\begin{aligned} & m\Delta x (\varphi_0^n)^\top (\pi^0 + \pi^-) P^{-1} Q u_0^n, \\ &= m\Delta x (\varphi_0^n)^\top (\pi^0 + \pi^-) K \tilde{g}^n + m\Delta x (\varphi_0^n)^\top (\pi^0 + \pi^-) K (\pi^0 + \pi^-) \varphi_0^n, \\ &= m\Delta x ((\pi^0 + \pi^-) \varphi_0^n)^\top ((\pi^0 + \pi^-) K \pi^+) \tilde{g}^n \\ &\quad + m\Delta x ((\pi^0 + \pi^-) \varphi_0^n)^\top ((\pi^0 + \pi^-) K (\pi^0 + \pi^-)) (\pi^0 + \pi^-) \varphi_0^n, \end{aligned}$$

and using the Young inequality leads to the following inequality

$$\begin{aligned}
& m\Delta x(\varphi_0^n)^\top(\pi^0 + \pi^-)P^{-1}Qu_0^n \\
& \leq \frac{m\Delta x}{2} \left(\|(\pi^0 + \pi^-)K\pi^+\| \left(|\tilde{g}^n|_{\mathbb{R}^p}^2 + |(\pi^0 + \pi^-)\varphi_0^n|_{\mathbb{R}^p}^2 \right) \right. \\
& \quad \left. + m\Delta x \left\| (\pi^0 + \pi^-)K(\pi^0 + \pi^-) \right\| |(\pi^0 + \pi^-)\varphi_0^n|_{\mathbb{R}^p}^2 \right).
\end{aligned}$$

Using the definition of $C_m'' \geq 0$ given by (3.11) in this last inequality enables to conclude. \square

4.2 Estimates for the discrete sequence in the domain

In this section, we derive estimates related to (3.2) that is used for the interior domain. In this regard, we introduce the following discrete Sobolev spaces

$$\begin{aligned}
\ell_{S_m}^2(\mathcal{I}) &= \left\{ v \in (\mathbb{R}^p)^\mathcal{I} \mid \sum_{i \in \mathcal{I}} v_i^\top S_m v_i \Delta x < +\infty \right\}, \\
h_{S_m}^1(\mathcal{I}) &= \left\{ v \in (\mathbb{R}^p)^\mathcal{I} \mid (v_i)_{i \in \mathcal{I}} \in \ell_{S_m}^2(\mathcal{I}), \quad ((v_{i+1} - v_i)/\Delta x)_{i \in \mathcal{I}} \in \ell_{S_m}^2(\mathcal{I}) \right\},
\end{aligned}$$

where $\mathcal{I} \subset \mathbb{N}$ and $S_m \in \mathcal{M}_{p,p}(\mathbb{R})$ is defined in (2.9). The norm induced by S_m in $\ell_{S_m}^2(\mathcal{I})$ is denoted

$$\|(v_i)_{i \in \mathcal{I}}\|_{\ell_{S_m}^2(\mathcal{I})}^2 = \sum_{i \in \mathcal{I}} v_i^\top S_m v_i \Delta x. \quad (4.6)$$

For sake of clarity, one denotes $\|(v_i)_{i \in \mathcal{I}}\|_{\ell_{S_m}^2(\mathcal{I})}^2 = \|(v_i)_{i \in \mathcal{I}}\|_{\ell_{S_m}^2}^2$ where the subset $\mathcal{I} \subset \mathbb{N}$ is implicitly given.

In order to ensure the convergence of all quantities introduced in the sequel, we consider $(u_i^n)_{i \geq 1} \in h_{S_m}^1(\mathbb{N}^*)$. This assumption is sufficient to enforce the convergence of the global quantities derived in the sequel and it is also consistent with compactly supported data usually used for practical applications. As a consequence, applying $\sum_{i \geq 1} (u_i^n)^\top S_m \cdot \Delta x$ to (3.2) we have

$$\begin{aligned}
& \sum_{i \geq 1} (u_i^n)^\top S_m \frac{u_i^{n+1} - u_i^n}{\Delta t} \Delta x \\
& = - \sum_{i \geq 1} (u_i^n)^\top S_m (\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n)) + \sum_{i \geq 1} (u_i^n)^\top S_m Q u_i^n \Delta x.
\end{aligned} \quad (4.7)$$

The following lemma consists in a first reformulation of the above equality.

Lemma 4.3 (Reformulation of (4.7)). *For all sequences $(u_i^n)_{i \geq 1} \in h_{S_m}^1(\mathbb{N}^*)$, (4.7) is equivalent to*

$$\begin{aligned}
& \frac{\|(u_i^{n+1})_{i \geq 1}\|_{\ell_{S_m}^2}^2 - \|(u_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2}{2\Delta t} \\
& = \frac{\Delta t}{2} \left(-(\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n))/\Delta x + Q u_i^n \right)_{i \geq 1} \|_{\ell_{S_m}^2}^2 + (u_1^n)^\top S_m \mathcal{F}(u_0^n, u_1^n) \\
& \quad + \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}(u_i^n, u_{i+1}^n) + \sum_{i \geq 1} (u_i^n)^\top S_m Q u_i^n \Delta x.
\end{aligned} \quad (4.8)$$

where the $\ell_{S_m}^2$ norm is defined in (4.6).

Proof. Since $S_m = S_m^\top$, the standard properties of the canonical inner product of \mathbb{R}^p associated to the scheme (3.2) yield to

$$\begin{aligned} & \sum_{i \geq 1} (u_i^n)^\top S_m \frac{u_i^{n+1} - u_i^n}{\Delta t} \Delta x \\ &= \frac{1}{2\Delta t} \sum_{i \geq 1} ((u_i^{n+1})^\top S_m u_i^{n+1} - (u_i^n)^\top S_m u_i^n) \Delta x - \frac{1}{2\Delta t} \sum_{i \geq 1} (u_i^{n+1} - u_i^n)^\top S_m (u_i^{n+1} - u_i^n) \Delta x \\ &= \frac{\|(u_i^{n+1})_{i \geq 1}\|_{\ell_{S_m}^2}^2 - \|(u_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2}{2\Delta t} - \frac{\Delta t}{2} \|(-(\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n))/\Delta x + Qu_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2. \end{aligned}$$

As a consequence, (4.7) is equivalent to

$$\begin{aligned} & \frac{\|(u_i^{n+1})_{i \geq 1}\|_{\ell_{S_m}^2}^2 - \|(u_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2}{2\Delta t} = \frac{\Delta t}{2} \|(-(\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n))/\Delta x + Qu_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2 \\ & - \sum_{i \geq 1} (u_i^n)^\top S_m (\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n)) + \sum_{i \geq 1} (u_i^n)^\top S_m Qu_i^n \Delta x. \end{aligned} \quad (4.9)$$

Now, using a discrete integration by parts for the second term on the right hand-side leads to

$$\begin{aligned} & - \sum_{i \geq 1} (u_i^n)^\top S_m (\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n)) \\ &= \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}(u_i^n, u_{i+1}^n) + (u_1^n)^\top S_m \mathcal{F}(u_0^n, u_1^n). \end{aligned}$$

Inserting this equality in (4.9) enables to get the expected reformulation. \square

We now give an upper bound for (4.8). For the convenient in the forthcoming developments, we now introduce the standard upwind numerical flux \mathcal{F}^{upw} defined as:

$$\mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) = \frac{1}{2} A(u_i^n + u_{i+1}^n) - \frac{1}{2} P|D|P^{-1}(u_{i+1}^n - u_i^n), \quad (4.10a)$$

$$= P\pi^+ DP^{-1} u_i^n + P(\pi^0 + \pi^-) DP^{-1} u_{i+1}^n. \quad (4.10b)$$

The above numerical flux can be seen as an example of a flux in (3.3).

Lemma 4.4. *For all sequences $(u_i^n)_{i \geq 1} \in h_{S_m}^1(\mathbb{N}^*)$, we have*

$$\begin{aligned} & \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}(u_i^n, u_{i+1}^n) + \sum_{i \geq 1} (u_i^n)^\top S_m Qu_i^n \Delta x \\ & \leq \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) + C_m'' \|(u_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2 \\ & \quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P(-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n), \end{aligned} \quad (4.11)$$

where $C_m'' \geq 0$ and $k(\cdot, \cdot) \in \mathcal{M}_{p,p}(\mathbb{R})$ are respectively defined in (3.11), (3.8).

Proof. In the one hand, according to the definition of $\mathcal{F}(u_i^n, u_{i+1}^n)$ given by (3.3) and to the consistency $\mathfrak{F}(u, u) = Au$, we have

$$\begin{aligned}
& \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}(u_i^n, u_{i+1}^n) \\
&= \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \left(\mathfrak{F}(u_i^n, u_{i+1}^n) - \frac{1}{2} (\mathfrak{F}(u_i^n, u_i^n) + \mathfrak{F}(u_{i+1}^n, u_{i+1}^n)) + \frac{1}{2} A(u_i^n + u_{i+1}^n) \right) \\
&\quad - \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P(\nu_{i+\frac{1}{2}}^n + |D|) P^{-1} (u_{i+1}^n - u_i^n), \\
&= \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) - \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P \nu_{i+\frac{1}{2}}^n P^{-1} (u_{i+1}^n - u_i^n) \\
&\quad + \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \left(\mathfrak{F}(u_i^n, u_{i+1}^n) - \frac{1}{2} (\mathfrak{F}(u_i^n, u_i^n) + \mathfrak{F}(u_{i+1}^n, u_{i+1}^n)) \right),
\end{aligned}$$

where we made appear the upwind numerical flux \mathcal{F}^{upw} (4.10a). Now, thanks to the Lipschitz continuity of \mathfrak{F} detailed in (3.8) we deduce

$$\begin{aligned}
& \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}(u_i^n, u_{i+1}^n) \\
&\leq \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) \\
&\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P(-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n).
\end{aligned} \tag{4.12}$$

On the other hand, performing the same computations as in (2.15) and (2.16) and using the definitions of $\gamma_m, C_m'' \geq 0$ given by (2.13), (3.11) we may write

$$\sum_{i \geq 1} (u_i^n)^\top S_m Q u_i^n \Delta x \leq \gamma_m \sum_{i \geq 1} (u_i^n)^\top S_m u_i^n \Delta x \leq C_m'' \| (u_i^n)_{i \geq 1} \|_{\ell_{S_m}^2}^2.$$

Adding this inequality with (4.12) enables to get (4.11). \square

According to Lemmata 4.3 and 4.4, we have the following:

$$\begin{aligned}
& \frac{\| (u_i^{n+1})_{i \geq 1} \|_{\ell_{S_m}^2}^2 - \| (u_i^n)_{i \geq 1} \|_{\ell_{S_m}^2}^2}{2\Delta t} \\
&\leq \frac{\Delta t}{2} \left\| \left(-(\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n)) / \Delta x + Q u_i^n \right)_{i \geq 1} \right\|_{\ell_{S_m}^2}^2 + (u_1^n)^\top S_m \mathcal{F}(u_0^n, u_1^n) \\
&\quad + \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) + C_m'' \| (u_i^n)_{i \geq 1} \|_{\ell_{S_m}^2}^2 \\
&\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P(-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n) I_p) P^{-1} (u_{i+1}^n - u_i^n).
\end{aligned} \tag{4.13}$$

4.3 Estimates for the discrete Riemann invariants in the boundary cell and for the discrete sequence in the domain

In this section, we associate the estimates reached in Subsections 4.1 and 4.2. Namely, considering (4.4) and (4.13), we now derive the following lemma:

Lemma 4.5. *Let us consider $(u_i^n)_{i \geq 1} \in h_{S_m}^1(\mathbb{N}^*)$ and let us denote $0 \leq R_m^n < +\infty$ such that*

$$\begin{aligned} R_m^n &= \left\| \left(-(\mathcal{F}(u_i^n, u_{i+1}^n) - \mathcal{F}(u_{i-1}^n, u_i^n)) / \Delta x + Qu_i^n \right)_{i \geq 1} \right\|_{\ell_{S_m}^2}^2 \\ &\quad + m \left| (\pi^0 + \pi^-) \left(-D \frac{\varphi_1^n - \varphi_0^n}{\Delta x} + P^{-1} Q u_0^n \right) \right|_{\mathbb{R}^p}^2. \end{aligned} \quad (4.14)$$

Considering $K \in \mathcal{M}_{p,p}(\mathbb{R})$, $C_m'' \geq 0$, $\mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n)$ are respectively defined in (2.6), (3.11), (4.10), \mathcal{E}^n defined in (3.7) satisfies

$$\begin{aligned} \frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{2\Delta t} &\leq \frac{\Delta t}{2} R_m^n + \frac{1}{2} (u_0^n)^\top S_m A u_0^n + (u_1^n)^\top S_m (\mathcal{F}(u_0^n, u_1^n) - \mathcal{F}^{\text{upw}}(u_0^n, u_1^n)) \\ &\quad - \frac{1}{2} \sum_{i \geq 0} (u_{i+1}^n - u_i^n)^\top S_m P |D| P^{-1} (u_{i+1}^n - u_i^n) \\ &\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P (-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n) \\ &\quad + \frac{m\Delta x}{2} \left\| (\pi^0 + \pi^-) K \pi^+ \right\| \|\tilde{g}^n\|_{\mathbb{R}^p}^2 + C_m'' \mathcal{E}^n. \end{aligned}$$

Proof. Summing (4.4) and (4.13), then using \mathcal{E}^n defined in (3.7) to reformulate $(\|(u_i^n)_{i \geq 1}\|_{\ell_{S_m}^2}^2, m\Delta x |(\pi^0 + \pi^-) \varphi_0^n|_{\mathbb{R}^p}^2)$, we deduce

$$\begin{aligned} &\frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{2\Delta t} \\ &\leq \frac{\Delta t}{2} R_m^n + \frac{m\Delta x}{2} \left\| (\pi^0 + \pi^-) K \pi^+ \right\| \|\tilde{g}^n\|_{\mathbb{R}^p}^2 + C_m'' \mathcal{E}^n \\ &\quad + \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) - m(\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) \\ &\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P (-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n) + (u_1^n)^\top S_m \mathcal{F}(u_0^n, u_1^n). \end{aligned} \quad (4.15)$$

The rest of the proof consists in rewriting (4.15). In this regard, from the definitions of $S_m, \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n)$ respectively given by (2.9), (4.10b), we deduce

$$\begin{aligned} S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) &= P^{-\top} (\pi^+ + m(\pi^0 + \pi^-)) P^{-1} \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n), \\ &= P^{-\top} (\pi^+ + m(\pi^0 + \pi^-)) P^{-1} (P \pi^+ D P^{-1} u_i^n + P(\pi^0 + \pi^-) D P^{-1} u_{i+1}^n), \\ &= P^{-\top} (\pi^+ + m(\pi^0 + \pi^-)) (\pi^+ D \varphi_i^n + (\pi^0 + \pi^-) D \varphi_{i+1}^n), \end{aligned}$$

where we recall $\varphi = P^{-1}u$. As a consequence, thanks to properties of the projectors $(\pi^\alpha)_{\alpha \in \{+,0,-\}}$ defined in (2.1), (4.15) reduces to

$$\begin{aligned} &\sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) - m(\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) \\ &= \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top \pi^+ D \varphi_i^n \\ &\quad + m \left(-(\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) + \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n \right). \end{aligned}$$

Now, Lemma A.1 is appealed to reformulate the above quantities that leads to

$$\begin{aligned}
& \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m \mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n) - m(\varphi_0^n)^\top (\pi^0 + \pi^-) D(\varphi_1^n - \varphi_0^n) \\
&= -(\varphi_1^n)^\top (\pi^+ D \varphi_0^n + m(\pi^0 + \pi^-) D \varphi_1^n) + \frac{1}{2} (\varphi_0^n)^\top (\pi^+ + m(\pi^0 + \pi^-)) D \varphi_0^n \\
&\quad - \frac{1}{2} \sum_{i \geq 0} (\varphi_{i+1}^n - \varphi_i^n)^\top (\pi^+ D - m(\pi^0 + \pi^-) D) (\varphi_{i+1}^n - \varphi_i^n) \\
&= - (u_1^n)^\top S_m \mathcal{F}^{\text{upw}}(u_0^n, u_1^n) + \frac{1}{2} (u_0^n)^\top S_m A u_0^n - \frac{1}{2} \sum_{i \geq 0} (u_{i+1}^n - u_i^n)^\top S_m P |D| P^{-1} (u_{i+1}^n - u_i^n),
\end{aligned}$$

where we used $S_m |D| = P^{-\top} (\pi^+ - m(\pi^0 + \pi^-)) P^{-1} D$. Inserting the above equality in (4.15), we deduce the expected upper-bound, that concludes the proof. \square

To conclude this subsection, we now give an upper-bound for $\frac{1}{2} (u_0^n)^\top S_m A u_0^n$. According to the definition of $u_0^n \in \mathbb{R}^p$ given by (3.4), this upper-bound is obtained in the same maner as in the proof of Proposition 2.5.

Lemma 4.6 (Upper-bound for $\frac{1}{2} (u_0^n)^\top S_m A u_0^n$). *Let assume Hypothesis 2.1. Let $m > 0$ and $S_m \in \mathcal{M}_{p,p}(\mathbb{R})$ under the form (2.9). For all $\alpha > 0$, we have*

$$\begin{aligned}
\frac{1}{2} (u_0^n)^\top S_m A u_0^n &\leq (1 + \alpha) \max_{i \in [1,p]} \max(0, \lambda_i) \|\pi^+ B \pi^+\|^2 |\tilde{g}^n|_{\mathbb{R}^p}^2 + (1 + \frac{1}{\alpha}) \|\pi^+ B \pi^0\|^2 |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 \\
&\quad + \left(\max_{i \in [1,p]} \max(0, \lambda_i) \|\pi^+ B \pi^-\|^2 + \frac{m}{2} \max_{i \in [1,p]} \min(0, \lambda_i) \right) |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2.
\end{aligned}$$

Proof. According to (2.9) and to the spectral decomposition $A = P D P^{-1}$, $\varphi = P^{-1} u$ we have

$$\begin{aligned}
(u_0^n)^\top S_m A u_0^n &= (\varphi_0^n)^\top \pi^+ D \varphi_0^n + m(\varphi_0^n)^\top \pi^0 D \varphi_0^n + m(\varphi_0^n)^\top \pi^- D \varphi_0^n, \\
&= (\varphi_0^n)^\top \pi^+ D \varphi_0^n + m(\varphi_0^n)^\top \pi^- D \varphi_0^n, \\
&\leq \max_{i \in [1,p]} \max(0, \lambda_i) |\pi^+ \varphi_0^n|_{\mathbb{R}^p}^2 + m \max_{i \in [1,p]} \min(0, \lambda_i) |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2. \tag{4.16}
\end{aligned}$$

Using (3.4) to reformulate $|\pi^+ \varphi_0^n|_{\mathbb{R}^p}^2$ then arguing the algebraic computations used in (2.22), we deduce

$$\begin{aligned}
|\pi^+ \varphi_0^n|_{\mathbb{R}^p}^2 &\leq 2(1 + \alpha) \|\pi^+ B \pi^+\|^2 |\tilde{g}^n|_{\mathbb{R}^p}^2 + 2(1 + \frac{1}{\alpha}) \|\pi^+ B \pi^0\|^2 |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 \\
&\quad + 2 \|\pi^+ B \pi^-\|^2 |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2.
\end{aligned}$$

Using this above estimate in (4.16), we deduce the expected result that concludes the proof. \square

4.4 Proof of Theorem 3.1

Proof. According to Lemma 4.5 and Lemma 4.6, we have

$$\begin{aligned}
\frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{2\Delta t} &\leq \frac{\Delta t}{2} R_m^n + (1 + \alpha) \max_{i \in [1, p]} \max(0, \lambda_i) \|\pi^+ B \pi^+\|^2 |\tilde{g}^n|_{\mathbb{R}^p}^2 + \left(1 + \frac{1}{\alpha}\right) \|\pi^+ B \pi^0\|^2 |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 \\
&\quad + \left(\max_{i \in [1, p]} \max(0, \lambda_i) \|\pi^+ B \pi^-\|^2 + \frac{m}{2} \max_{i \in [1, p]} \min(0, \lambda_i) \right) |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2 \\
&\quad + (u_1^n)^\top S_m (\mathcal{F}(u_0^n, u_1^n) - \mathcal{F}^{\text{upw}}(u_0^n, u_1^n)) \\
&\quad - \frac{1}{2} \sum_{i \geq 0} (u_{i+1}^n - u_i^n)^\top S_m P |D| P^{-1} (u_{i+1}^n - u_i^n) \\
&\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P (-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n) \\
&\quad + \frac{m\Delta x}{2} \|\pi^0 + \pi^-\| K \pi^+ \|\pi^+\|^2 |\tilde{g}^n|_{\mathbb{R}^p}^2 + C_m'' \mathcal{E}^n,
\end{aligned}$$

where $K \in \mathcal{M}_{p,p}(\mathbb{R})$, $C_m'' \geq 0$, $\mathcal{F}^{\text{upw}}(u_i^n, u_{i+1}^n)$ are respectively defined in (2.6), (3.11), (4.10). Adding the estimate derived in Lemma 4.1 to the above inequality and using the definition of $C_\alpha, C'_\alpha > 0$, $\text{Kr}_m \in \mathbb{R}$ given in (2.18), we deduce

$$\begin{aligned}
&\frac{|\pi^0 \varphi_0^{n+1}|_{\mathbb{R}^p}^2 - |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2}{2\Delta t} + \frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{2\Delta t} \\
&\leq \left(C_\alpha + \frac{m\Delta x}{2} \|\pi^0 + \pi^-\| K \pi^+ \|\pi^+\| \right) |\tilde{g}^n|_{\mathbb{R}^p}^2 + C'_\alpha |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + C_m'' \mathcal{E}^n - \text{Kr}_m |\pi^- \varphi_0^n|_{\mathbb{R}^p}^2 \\
&\quad + \frac{\Delta t}{2} R_m^n + (u_1^n)^\top S_m (\mathcal{F}(u_0^n, u_1^n) - \mathcal{F}^{\text{upw}}(u_0^n, u_1^n)) \\
&\quad - \frac{1}{2} \sum_{i \geq 0} (u_{i+1}^n - u_i^n)^\top S_m P |D| P^{-1} (u_{i+1}^n - u_i^n) \\
&\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P (-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n).
\end{aligned}$$

Since $m > 0$ is such that $\text{Kr}_m \geq 0$, the proof of Theorem 3.1 will be completed as soon as the following inequality is satisfied:

$$\begin{aligned}
&\frac{\Delta t}{2} R_m^n + (u_1^n)^\top S_m (\mathcal{F}(u_0^n, u_1^n) - \mathcal{F}^{\text{upw}}(u_0^n, u_1^n)) \\
&\quad - \frac{1}{2} \sum_{i \geq 0} (u_{i+1}^n - u_i^n)^\top S_m P |D| P^{-1} (u_{i+1}^n - u_i^n) \\
&\quad + \frac{1}{2} \sum_{i \geq 1} (u_{i+1}^n - u_i^n)^\top S_m P (-\nu_{i+\frac{1}{2}}^n + k(u_i^n, u_{i+1}^n)) P^{-1} (u_{i+1}^n - u_i^n) \leq 0,
\end{aligned} \tag{4.17}$$

where R_m^n is defined in (4.14). But, since

$$\mathcal{F}(u_0^n, u_1^n) - \mathcal{F}^{\text{upw}}(u_0^n, u_1^n) = \mathfrak{F}(u_0^n, u_1^n) - \frac{1}{2} A(u_1^n + u_0^n) - P \nu_{\frac{1}{2}}^n P^{-1} (u_1^n - u_0^n),$$

and $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ is such that (3.9) holds, we deduce the existence of $\Delta t > 0$ such that (4.17) holds that concludes the proof. \square

To conclude this section, we summarize the procedure presented in Section 3. For a given IBVP under the form (1.1) and a given numerical flux satisfying (3.8) and $\mathfrak{F}(u, u) = Au$, the procedure is the following:

- First, we use Lemma 2.2 to determine $B \in \mathcal{M}_{p,p}(\mathbb{R})$ so that (1.1) can be reformulated under the form (2.5).
- Second, considering $S_m \in \mathcal{M}_{p,p}(\mathbb{R})$ as defined in (2.9), $m \geq 0$ has to enforce $\text{Kr}_m \geq 0$ in (2.18c). Such a choice enables to define (3.10) as a criteria for the discrete stability.
- Third, according to Theorem 3.1, (3.10) is reachable if the sequence $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ is such that (3.13) and (3.14) hold.
- Finally, $\Delta t > 0$ has to be selected in such a way (4.17) (or equivalently (3.10)) holds. When $m > 0$ increases in (2.9) then R_m^n defined in (4.14) also increases and thus the time step $\Delta t > 0$ decreases. In practice, m will be chosen to minimize R_m^n under the constraint $\text{Kr}_m \geq 0$.

The next section gives practical applications of this procedure in numerical experiments.

5 Numerical experiments

This section is dedicated to the implementation of the numerical strategy presented before. First of all, a practical choice of $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ is proposed for some well-known numerical fluxes (Lax-Wendroff or HLL) and the resulting CFL condition. Then, two numerical tests are investigated (wave equation and relaxed dispersive wave model) and comparisons are performed with classical approaches to emphasize the benefit of our method.

5.1 Viscous correction for standard numerical fluxes

This section gives practical choices of the viscous correction procedure (3.3) for specific, but commonly used numerical flux $\mathfrak{F}(u_i^n, u_{i+1}^n)$. Namely, we consider the following form :

$$\mathfrak{F}(u_i^n, u_{i+1}^n) = \frac{1}{2}A(u_i^n + u_{i+1}^n) - \frac{1}{2}P\Lambda P^{-1}(u_{i+1}^n - u_i^n), \quad (5.1)$$

where Λ is a diagonal semi-definite positive matrix in $\mathcal{M}_{p,p}(\mathbb{R})$. Several well-known numerical fluxes take the form (5.1). For instance, the Lax-Wendroff numerical flux corresponds to $\Lambda = \frac{\Delta t}{\Delta x} D^2$ and the symmetric HLL (also called Rusanov) numerical flux is such that $\Lambda = \rho(|D|)I_p$ where $\rho(|D|) > 0$ denotes the spectral radius of $|D|$. Such numerical fluxes are obviously consistent and they satisfy the k -Lipschitz continuity assumption (3.8) with $k(u_i^n, u_{i+1}^n) = -\Lambda$.

Considering (5.1), and according to Theorem 3.1, $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ has to be selected in a such way (3.13) and (3.14) hold. Since $\Lambda \in \mathcal{M}_{p,p}(\mathbb{R})$ in (5.1) is a diagonal semi-definite positive matrix, (3.14) is satisfied for the choice

$$\nu_{i+\frac{1}{2}}^n = 0, \quad \forall i \geq 1. \quad (5.2)$$

We now have to deal with $\nu_{\frac{1}{2}}^n$ to enforce (3.13). In the case of (5.1), (3.13) is equivalent to

$$-\frac{1}{2}(u_1^n)^\top S_m P (\Lambda + \nu_{\frac{1}{2}}^n) P^{-1} (u_1^n - u_0^n) \leq 0.$$

As a consequence, we use the following admissible choice:

$$\nu_{\frac{1}{2}}^n = \begin{cases} 0, & \text{if } (u_1^n)^\top S_m P \Lambda P^{-1} (u_1^n - u_0^n) \geq 0, \\ -\Lambda, & \text{otherwise.} \end{cases} \quad (5.3)$$

The above formulation defines a non-linear limitation procedure to enforce (3.13). Equipped with this procedure, the time step $\Delta t > 0$ has to be chosen according to (4.17). However, (4.17) is defined by global quantities that are difficult to use for the practical implementations. As a consequence, we propose the following heuristic CFL-like condition:

$$\frac{\Delta t}{\Delta x} (|D| + \max_i |\nu_{i+\frac{1}{2}}^n|) \leq \frac{1}{2}, \quad (5.4)$$

that is inspired from [9, 25]. We now illustrate the above choices on two test cases.

5.2 Example of weak numerical instabilities

This section gives an example of weak instabilities and the relevancy of the viscous correction (3.3) for the stabilization. In this regard, we consider the wave equations that governs $u = (u_1, u_2)^\top \in \mathbb{R}^2$ as follows:

$$\begin{cases} \partial_t u + A \partial_x u = 0, & \forall (t, x) \in (\mathbb{R}_*^+)^2, \\ \mathbf{N}u(t, 0) = g(t), & \forall t \geq 0, \\ u(t = 0, x) = u^{\text{in}}(x), & \forall x > 0, \end{cases} \quad (5.5)$$

where the matrices $(A, \mathbf{N}) \in \mathcal{M}_{2,2}(\mathbb{R}) \times \mathcal{M}_{1,2}(\mathbb{R})$ are given by

$$A = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}, \quad a > 0, \quad \mathbf{N} = \begin{pmatrix} 1 & 0 \end{pmatrix}. \quad (5.6)$$

According to the above definitions, there is no null eigenvalue in A (*i.e.* $\pi^0 = 0$) so $\alpha = 0$ in (2.18) is selected. In addition, the matrix $B \in \mathcal{M}_{2,2}(\mathbb{R})$ derived for (2.5) and a suitable symmetrizer $S_m \in \mathcal{M}_{2,2}(\mathbb{R})$ are given by

$$B = \begin{pmatrix} \sqrt{2} & 1 \\ 0 & 1 \end{pmatrix}, \quad S_2 = \frac{1}{2} \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

The above formulation of S_2 corresponds to $m = 2$ in (2.9) and $\text{Kr}_m = 0$ in (2.18c). Let us remark that this choice of m is such that R_m^n defined in (4.14) is minimal. As a consequence, straightforward computation shows that (3.10) is equivalent $\mathcal{S}^n \leq 0$ where $\mathcal{S}^n \in \mathbb{R}$ is defined as follows:

$$\mathcal{S}^n := \frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{2\Delta t} - 2a |\tilde{g}^n|_{\mathbb{R}^p}^2. \quad (5.7)$$

In order to check $\mathcal{S}^n \leq 0$ with the choices given in Subsection 5.1, the space domain $[0, 15]$ is discretized with 2 000 cells and we consider the following initial and boundary conditions:

$$u^{\text{in}}(x) = e^{-2x} (\sin(2\pi x), \cos(2\pi x))^\top, \quad g(t) = \sin(2\pi t). \quad (5.8)$$

Homogeneous Neumann boundary conditions are prescribed on the right side of the domain. According to (5.8) and since the computational domain is sufficiently large, this discrete boundary does not influence the numerical simulation. On the left side, the discrete boundary conditions are prescribed by $u_0^n \in \mathbb{R}^2$ defined in (3.4). We consider the Lax-Wendroff (LW) numerical flux in (5.1) then $(\nu_{i+\frac{1}{2}}^n)_{i \geq 0}$ and $\Delta t > 0$ are selected according to (5.2), (5.3), (5.4). We also consider a classical approach based on the approximation of (2.5) with the scheme (1.3), (3.4), (3.6) in which the numerical flux $\mathfrak{F}(u_i^n, u_{i+1}^n)$ is chosen as the LW flux. Hence, the comparison between the two approaches will enable to emphasize the importance of the artificial viscosity in (3.3).

Figure 2 shows the time evolution of \mathcal{S}^n defined in (5.7) and also the space dependency of the solution $x \in [0, 0.2] \mapsto u_i(t = 0.6, x)$, $i = 1, 2$.

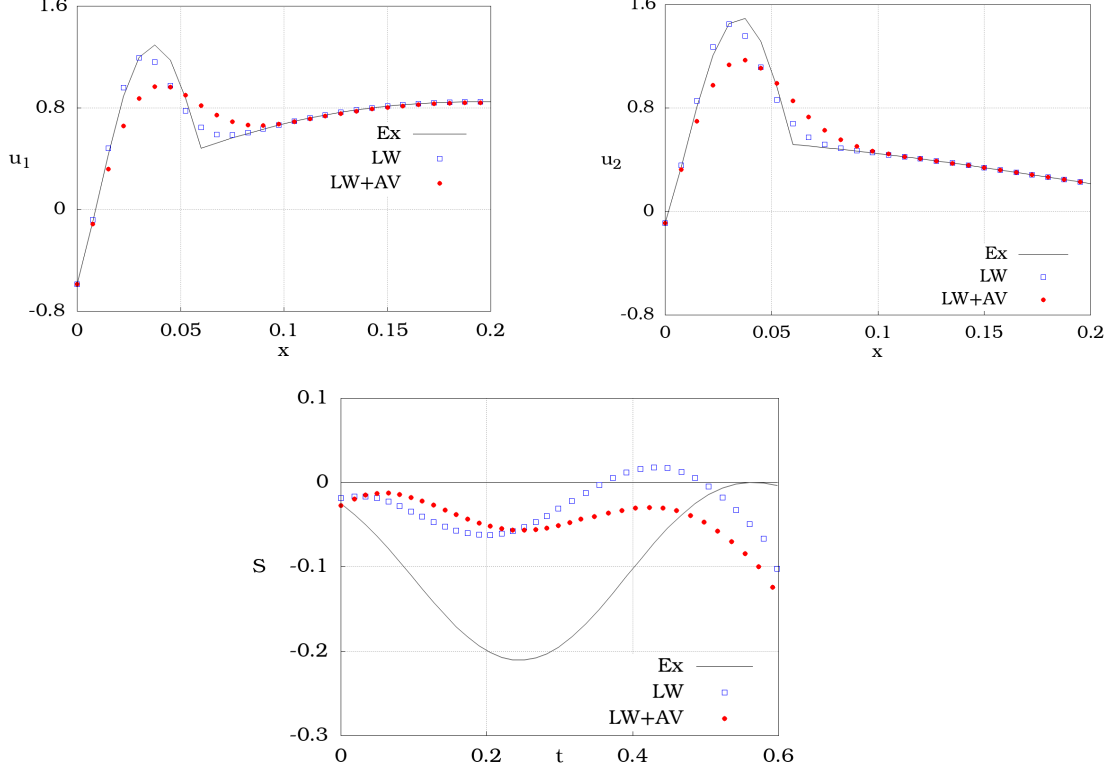


Figure 2: Top: $x \in [0, 0.2] \mapsto u_i(t = 0.6, x)$ for the problem (5.5)-(5.8). Bottom: time evolution of \mathcal{S}^n defined in (5.7). The legend is: LW corresponds to the Lax-Wendroff scheme (1.3), LW+AV corresponds to the viscous Lax-Wendroff scheme (3.2) (both schemes are endowed with the discrete boundary conditions (3.4)-(3.6) and Ex corresponds to the exact solution).

The standard Lax-Wendroff scheme (1.3) and its viscous version (3.2) are both consistent but the standard version becomes unstable close to $t = 0.4$. The viscous correction as defined in Subsection 5.1 ensures $\mathcal{S}^n \leq 0$ but the numerical results for u_1 and u_2 are more diffusive.

5.3 Relaxed dispersive waves model

The second numerical experiment is devoted to the relaxation of a dispersive wave model. Namely, we consider $u = (u_1, u_2, v, \partial_t v)^\top \in \mathbb{R}^4$ governed as follows:

$$\begin{cases} \partial_t u + A \partial_x u = Qu, & \forall (t, x) \in (\mathbb{R}_*^+)^2, \\ Nu(t, 0) = g(t), & \forall t \geq 0, \\ u(t = 0, x) = u^{\text{in}}(x), & \forall x > 0, \end{cases} \quad (5.9)$$

where the matrices $(A, Q, N) \in (\mathcal{M}_{4,4}(\mathbb{R}))^2 \times \mathcal{M}_{1,4}(\mathbb{R})$ are given by

$$A = \begin{pmatrix} 0 & a & 0 & -1 \\ a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{\kappa^2}{\varepsilon^2} & 0 & 0 & 0 \end{pmatrix}, \quad a, \kappa, \varepsilon > 0, \quad Q = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{\varepsilon^2} & 0 \end{pmatrix}, \quad N = (1 \ 0 \ 0 \ 0). \quad (5.10)$$

This system can be derived from dispersive Boussinesq type equations (5.13) presented below. The above A matrix is a \mathbb{R} -diagonalisable matrix and its eigenvalues are

$(\sqrt{a^2 + (\kappa/\varepsilon)^2}, 0, 0, -\sqrt{a^2 + (\kappa/\varepsilon)^2})$. The initial and boundary conditions u^{in}, g are

$$u^{\text{in}}(x) = \left(1, -\frac{a\tau}{\ell}, -\frac{(\kappa\tau)^2}{\ell(\varepsilon^2 + \tau^2)}, \frac{\kappa^2\tau}{\ell(\varepsilon^2 + \tau^2)}\right)^\top e^{-\frac{x}{\ell}}, \quad g(t) = e^{-\frac{t}{\tau}}, \quad (5.11)$$

where $\tau > 0$ is a free parameter, and $\ell = \sqrt{(a\tau)^2 + (\kappa\tau)^2/(\varepsilon^2 + \tau^2)}$. Under the above conditions, the exact $H^1((\mathbb{R}_*^+)^2, \mathbb{R}^4)$ solution (5.9) is $u(t, x) = u^{\text{in}}(x)g(t)$ where u^{in} and g are given by (5.11). In addition, it can easily be checked

$$\forall(t, x) \in (\mathbb{R}_*^+)^2, \quad (u_1, u_2, v, \partial_t v)(t, x) \xrightarrow{\varepsilon \rightarrow 0} (w_1, w_2, \kappa^2 \partial_x w_1, \kappa^2 \partial_{tx}^2 w_1)(t, x), \quad (5.12)$$

where $(w_1, w_2)^\top \in \mathbb{R}^2$ is the solution of the dispersive problem (5.13). Hence, (5.9) formally stands for an ε -relaxed version of the following dispersive wave model

$$\begin{cases} \begin{pmatrix} 1 - \kappa^2 \partial_x^2 & 0 \\ 0 & 1 \end{pmatrix} \partial_t \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} \partial_x \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = 0, & \forall(t, x) \in (\mathbb{R}_*^+)^2, \\ w_1(t, 0) = g(t), & \forall t \geq 0, \\ w(t = 0, x) = w^{\text{in}}(x), & \forall x > 0, \end{cases} \quad (5.13)$$

where $w = (w_1, w_2)^\top \in \mathbb{R}^2$, $a > 0$ and $g(t) \in \mathbb{R}$. The above dispersive problem corresponds to the formal limit $\varepsilon \rightarrow 0$ in (5.9). Equation (5.13) is called linearized Boussinesq-Abbott equation. The solution w_2 of (5.13) is good approximation of surface elevation in linearized shallow water regime. The dispersion in (5.13) induces the computation of non-local terms. This is why several works transform this equation into an hyperbolic problem by relaxation. Note that, the relaxation introduces zero eigenvalues (see [20, 21, 36]). Moreover, for $\tau > 0$ and considering the following initial and boundary conditions

$$w^{\text{in}}(x) = (1, -a\tau/\sqrt{(a\tau)^2 + \kappa^2})^\top e^{-\frac{x}{\sqrt{(a\tau)^2 + \kappa^2}}}, \quad g(t) = e^{-\frac{t}{\tau}}, \quad (5.14)$$

the solution w to (5.13)-(5.14) belongs to $H^1((\mathbb{R}_*^+)^2, \mathbb{R}^2)$ and is given by

$$w(t, x) = w^{\text{in}}(x)g(t). \quad (5.15)$$

Now, we have to compute $B \in \mathcal{M}_{4,4}(\mathbb{R})$ for the boundary conditions in (3.4). According to (5.9), straightforward computations lead to

$$B = \begin{pmatrix} \frac{1}{\sqrt{a^2 + (\kappa/\varepsilon)^2}} & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The second step consists in defining $S_m \in \mathcal{M}_{4,4}(\mathbb{R})$ under the form (2.9). Using the above definition of B , we select $m > 0$ in (2.9) such that $\text{Kr}_m = 0$ in (2.18c). In addition, since straightforward computations shows that $\pi^+ B \pi^0 = 0$, we fix $\alpha = 0$ in (2.18). As a consequence, (3.10) is equivalent to $\mathcal{S}^n \leq 0$, where \mathcal{S}^n is defined as follows:

$$\begin{aligned} \mathcal{S}^n := & \frac{(|\pi^0 \varphi_0^{n+1}|_{\mathbb{R}^p}^2 + \mathcal{E}^{n+1}) - (|\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 + \mathcal{E}^n)}{2\Delta t} \\ & - (C_\alpha + \frac{m\Delta x}{2} \|(\pi^0 + \pi^-)K\pi^+\|) |\tilde{g}^n|_{\mathbb{R}^p}^2 - C'_\alpha |\pi^0 \varphi_0^n|_{\mathbb{R}^p}^2 - C''_m \mathcal{E}^n, \end{aligned} \quad (5.16)$$

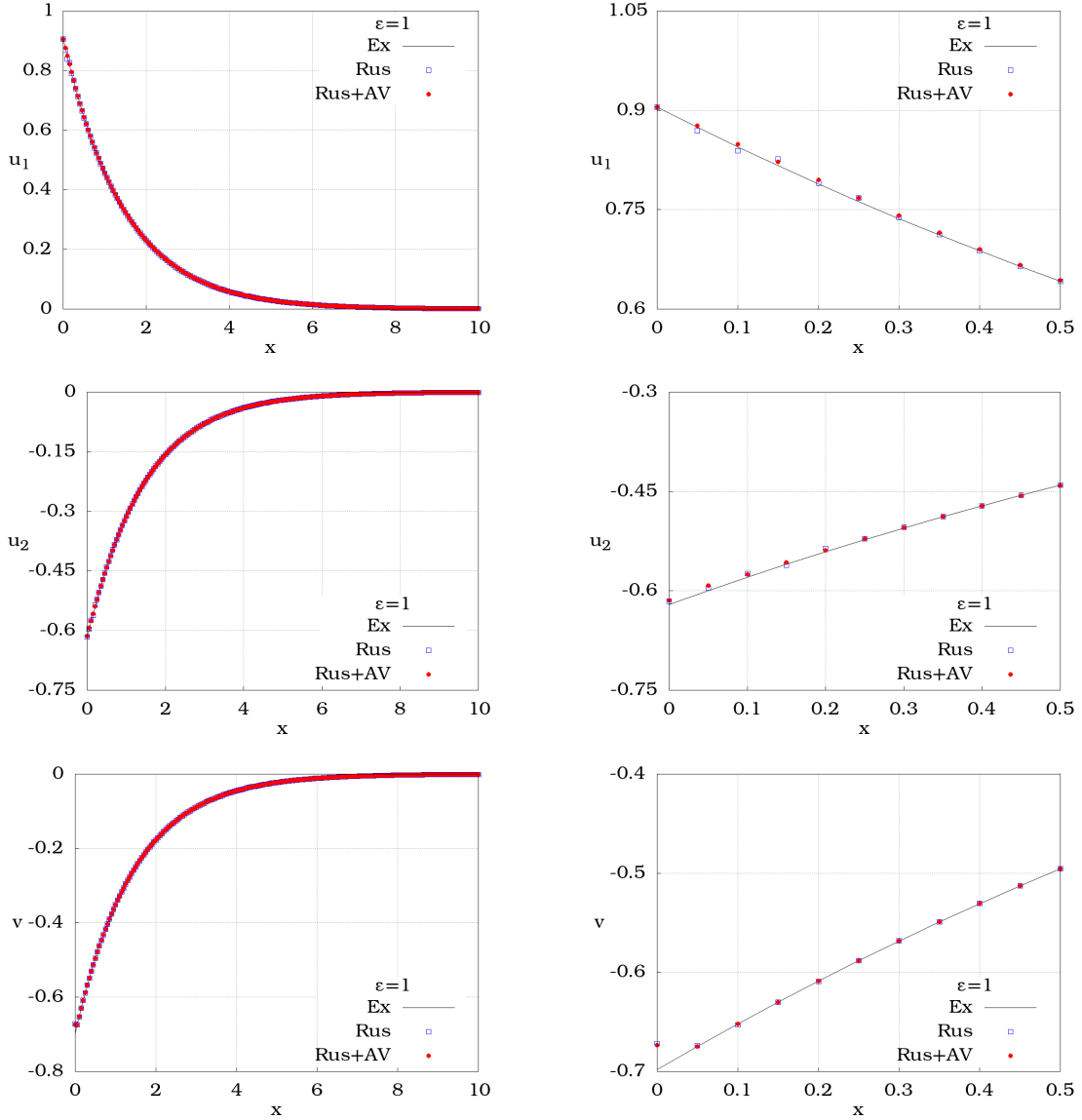
where $K \in \mathcal{M}_{4,4}(\mathbb{R})$ and the constants $C_\alpha, C'_\alpha, C''_m > 0$ are respectively given by (2.6), (2.18) (3.11) (note that they can be explicitly computed in this case). Even if the inequality $\mathcal{S}^n \leq 0$

stands for a criteria to assess the numerical stability of the schemes, it is no longer relevant in the regimes $\varepsilon \rightarrow 0$ because $C_m'' \underset{\varepsilon \rightarrow 0}{\sim} \frac{1}{\varepsilon^2}$ that leads to

$$\lim_{\varepsilon \rightarrow 0} C_m'' = +\infty. \quad (5.17)$$

For the first numerical simulation, (5.9)-(5.11) are endowed with $(a, \tau, \kappa, \varepsilon)^\top = (1, 1, 1.5, 1)^\top$ and we consider the spatial domain $[0, 20]$ discretized with 400 cells. Homogeneous Neumann boundary conditions are prescribed on the right side of the domain. On the left side, the discrete boundary conditions are given by $u_0^n \in \mathbb{R}^4$ defined in (3.4). We consider the Rusanov (Rus) numerical flux in (5.1) then $(v_{i+\frac{1}{2}}^n)_{i \geq 0}$ and $\Delta t > 0$ are selected according to (5.2), (5.3), (5.4).

Figure 3 reports the exact solution (labeled as Ex) $x \mapsto (u_1, u_2, v, \partial_t v)(t = 0.1, x) = u^{\text{in}}(x)g(t = 0.1)$ where $u^{\text{in}}(x)$ and $g(t)$ are given by (5.11), and the numerical solutions computed by the standard solver Rus and the viscous one Rus+AV. One can observe the numerical solutions are in a very good agreement with the exact solution. In Figure 4, we plot the time evolution of \mathcal{S}^n and we observe the two schemes are stable in the sense $\mathcal{S}^n \leq 0$.



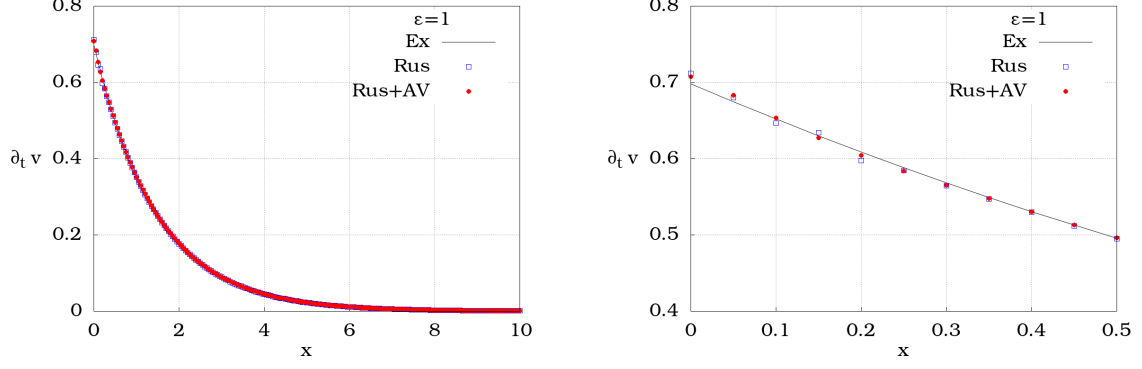


Figure 3: Relaxed dispersive model ($\varepsilon = 1$). Left: $x \in [0, 10] \mapsto (u_1, u_2, v, \partial_t v)(t = 0.1, x)$. Right: $x \in [0, 0.5] \mapsto (u_1, u_2, v, \partial_t v)(t = 0.1, x)$. The legend is: Ex corresponds to the exact solution, Rus corresponds to the Rusanov scheme (1.3), Rus+AV corresponds to the viscous Rusanov scheme (3.2) (both schemes are endowed with the discrete boundary conditions (3.4), (3.6)) and Ex corresponds to the exact solution).

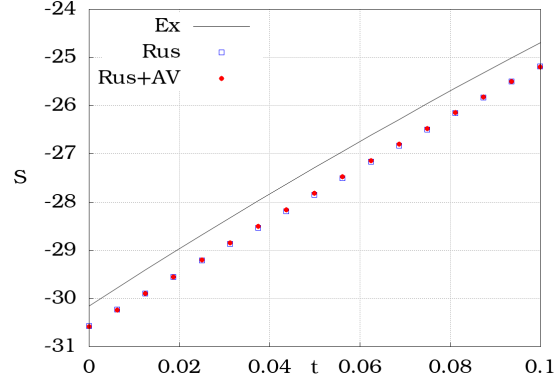


Figure 4: Relaxed dispersive model ($\varepsilon = 1$). Time evolution of \mathcal{S}^n defined in (5.16). The legend is: Ex corresponds to the exact solution, Rus corresponds to the Rusanov scheme (1.3), Rus+AV corresponds to the viscous Rusanov scheme (3.2) (both schemes are endowed with the discrete boundary conditions (3.4), (3.6)) and Ex corresponds to the exact solution).

In a second test, we numerically investigate the convergence of the relaxation system (5.9) towards the dispersive equations (5.13). In this regard, we consider $(a, \tau, \kappa, \varepsilon)^\top = (1, 1, 1.5, 10^{-2})^\top$ in (5.9),(5.11). The numerical framework is the same as before except the space domain $[0, 20]$ which is discretized with 400 cells. Figure 5 reports the solutions at time $t = 0.1$ obtained by the standard method, the viscous one and the asymptotic solution given by (5.15). According to (5.12), the components of the exact solution of the dispersive model is thus $\tilde{w}(t, x) = (w_1, w_2, \kappa^2 \partial_x w_1, \kappa^2 \partial_{tx}^2 w_1)(t, x)$ where $w(t, x) = (w_1, w_2)(t, x)$ is given by (5.15)-(5.14). For clarity, the plots of the four components of the exact solution \tilde{w} and the numerical solutions are only reported for $x \in [0, 10]$. Figure 6 reports the time evolution of \mathcal{S}^n defined by (5.16) obtained by the exact solution of the relaxed dispersive model (5.9) with $\varepsilon = 10^{-2}$ and obtained by the two numerical methods (Rus and Rus+AV). In this case, we observe the evolution of \mathcal{S}^n has no longer meaning in vertu of (5.17). However, one can observe the standard Rusanov scheme (1.3) coupled to the discrete boundary conditions (3.4)-(3.6) clearly fails. Regarding the viscous version (3.2) endowed with our non-linear limitation procedure (5.3), in spite of some errors localized at the boundary, the method is genuinely more stable (see Figure 6).

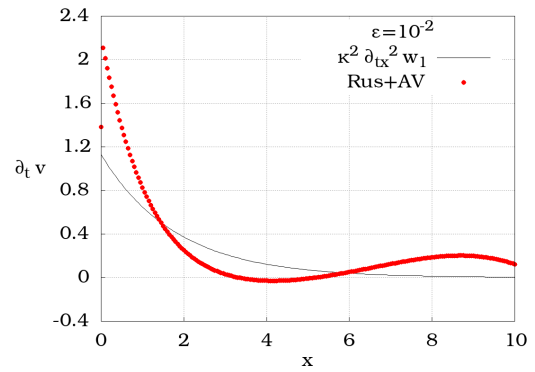
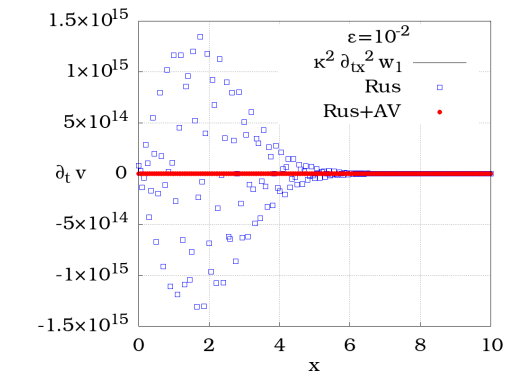
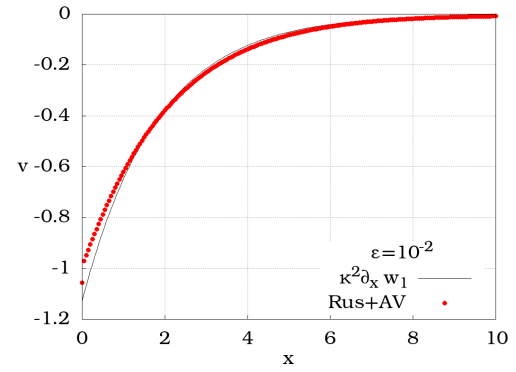
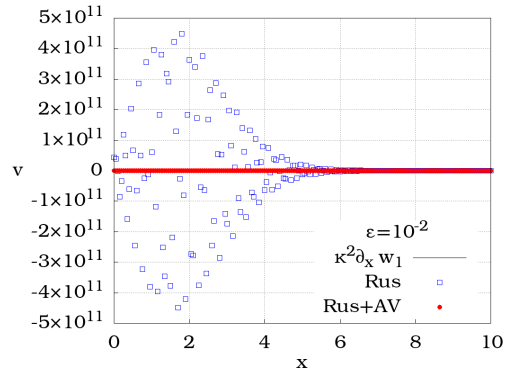
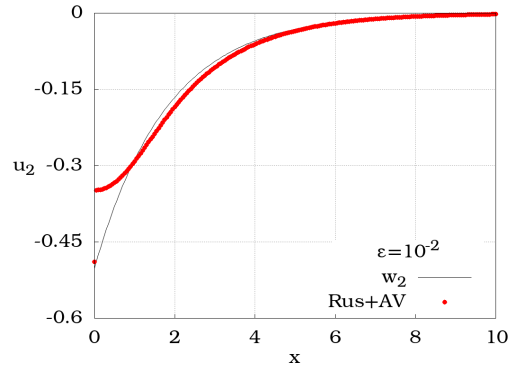
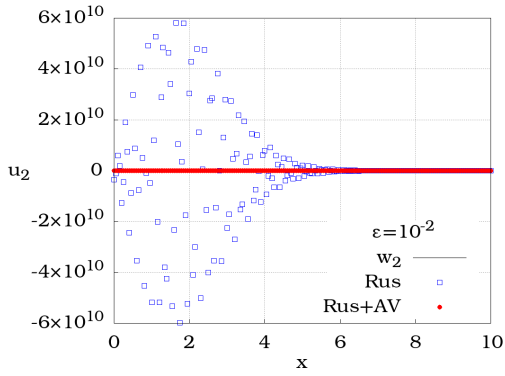
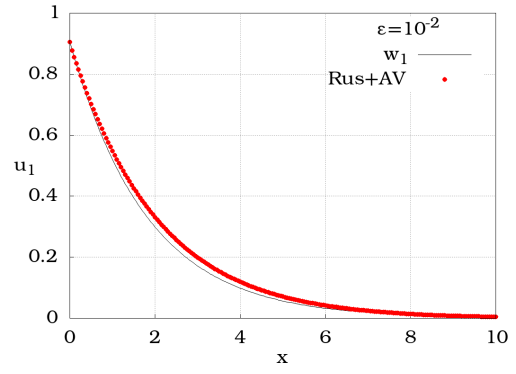
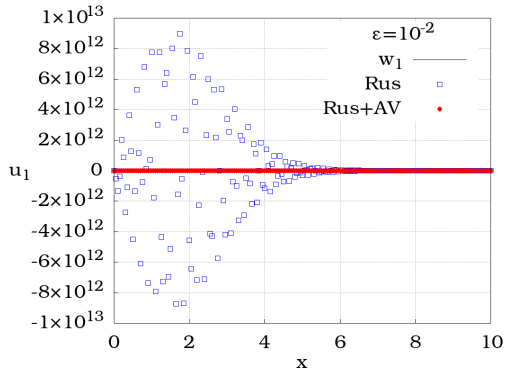


Figure 5: Relaxed dispersive model ($\varepsilon = 10^{-2}$). Left: $x \in [0, 10] \mapsto (u_1, u_2, v, \partial_t v)(t = 0.1, x)$. Right: $x \in [0, 10] \mapsto (u_1, u_2, v, \partial_t v)(t = 0.1, x)$ (only the viscous and the exact asymptotic solutions $\tilde{w} = (w_1, w_2, \kappa^2 \partial_x w_1, \kappa^2 \partial_{tx}^2 w_1)(t, x)$ where $w(t, x) = (w_1, w_2)(t, x)$ is given by (5.15)-(5.14)). The legend is: Rus corresponds to the Rusanov scheme (1.3), Rus+AV corresponds to the viscous Rusanov scheme (3.2) (both schemes are endowed with the discrete boundary conditions (3.4), (3.6)) and Ex corresponds to the exact solution.

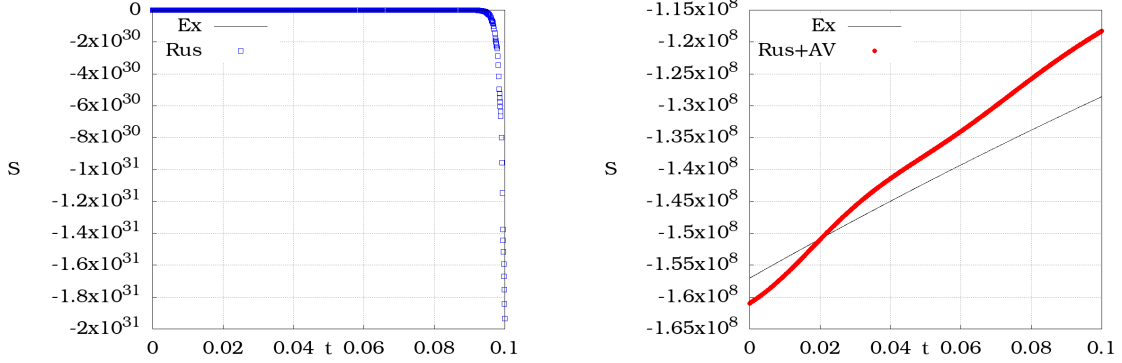


Figure 6: Relaxed dispersive model ($\varepsilon = 10^{-2}$). Time evolution of \mathcal{S}^n defined in (5.16). Left: exact and standard Rusanov solutions. Right: exact and viscous solutions.

6 Conclusion

We have proposed an explicit procedure to correct three-point finite volume schemes devoted to the numerical approximation of linear hyperbolic initial boundary value problems. The fully discrete stability of the corrected scheme is proved under restrictions for the artificial viscosity and for the time step. These numerical parameters are selected with heuristic procedures in the numerical experiments that concern the standard hyperbolic wave model and a relaxed dispersive model. The numerical results obtained on these two test cases confirm the discrete stability estimates. This work is a first step toward the investigation of fully discrete stability of wave-structure interactions problems (see [3], [4], [5] and [6]). In that case, the linear hyperbolic regime is not necessarily a good regime. Indeed, this regime omits the dispersive and non-linear effects that play an important role in coastal areas (see [30]). In next works, we want to generalize our strategy to non-linear and dispersive problem (such as the one in (5.13)) involving boundary conditions.

A Technical lemma

Lemma A.1. *Let us consider $(\varphi_i^n)_{i \geq 0} \subset \mathbb{R}^p$ a compactly supported sequence such that*

$$\sum_{i \geq 0} |\varphi_i^n|_{\mathbb{R}^p}^2 \Delta x < +\infty, \quad \sum_{i \geq 0} |(\varphi_{i+1}^n - \varphi_i^n) / \Delta x|_{\mathbb{R}^p}^2 \Delta x < +\infty.$$

Let also us consider $D \in \mathcal{M}_{p,p}(\mathbb{R})$ the diagonal matrix that contains the eigenvalues of A organized in descending order and also the diagonal projectors $(\pi^\alpha)_{\{+,0,-\}}$ given by (2.1). The

following reformulations hold:

$$\begin{aligned}
(i) \quad & \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top \pi^+ D \varphi_i^n \\
& = -(\varphi_1^n)^\top \pi^+ D \varphi_0^n + \frac{1}{2} (\varphi_0^n)^\top \pi^+ D \varphi_0^n - \frac{1}{2} \sum_{i \geq 0} (\varphi_{i+1}^n - \varphi_i^n)^\top \pi^+ D (\varphi_{i+1}^n - \varphi_i^n), \\
(ii) \quad & -(\varphi_0^n)^\top (\pi^0 + \pi^-) D (\varphi_1^n - \varphi_0^n) + \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n \\
& = -(\varphi_1^n)^\top (\pi^0 + \pi^-) D \varphi_1^n + \frac{1}{2} (\varphi_0^n)^\top (\pi^0 + \pi^-) D \varphi_0^n \\
& \quad + \frac{1}{2} \sum_{i \geq 0} (\varphi_{i+1}^n - \varphi_i^n)^\top (\pi^0 + \pi^-) D (\varphi_{i+1}^n - \varphi_i^n).
\end{aligned}$$

Proof. Concerning (i), using the standard indices translation, a direct computation gives

$$\begin{aligned}
& \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top \pi^+ D \varphi_i^n \\
= & \sum_{i \geq 2} (\varphi_i^n)^\top \pi^+ D \varphi_{i-1}^n - \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_i^n \\
= & -(\varphi_1^n)^\top \pi^+ D \varphi_0^n + \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_{i-1}^n - \left(\frac{1}{2} + \frac{1}{2} \right) \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_i^n \\
= & -(\varphi_1^n)^\top \pi^+ D \varphi_0^n - \frac{1}{2} \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_i^n + \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_{i-1}^n - \frac{1}{2} \sum_{i \geq 2} (\varphi_{i-1}^n)^\top \pi^+ D \varphi_{i-1}^n.
\end{aligned}$$

Reorganizing the terms, the above formulation writes

$$\begin{aligned}
& \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top \pi^+ D \varphi_i^n \\
= & -(\varphi_1^n)^\top \pi^+ D \varphi_0^n + \frac{1}{2} (\varphi_0^n)^\top \pi^+ D \varphi_0^n \\
& \quad - \frac{1}{2} \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_i^n + \sum_{i \geq 1} (\varphi_i^n)^\top \pi^+ D \varphi_{i-1}^n - \frac{1}{2} \left(\sum_{i \geq 2} (\varphi_{i-1}^n)^\top \pi^+ D \varphi_{i-1}^n + \varphi_0^n \pi^+ D \varphi_0^n \right), \\
= & -(\varphi_1^n)^\top \pi^+ D \varphi_0^n + \frac{1}{2} (\varphi_0^n)^\top \pi^+ D \varphi_0^n - \frac{1}{2} \sum_{i \geq 1} (\varphi_i^n - \varphi_{i-1}^n)^\top \pi^+ D (\varphi_i^n - \varphi_{i-1}^n),
\end{aligned}$$

that is the expected reformulation.

For (ii), with direct straightforward computations, we have

$$\begin{aligned}
& -(\varphi_0^n)^\top (\pi^0 + \pi^-) D (\varphi_1^n - \varphi_0^n) + \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n \\
= & (\varphi_0^n)^\top (\pi^0 + \pi^-) D \varphi_0^n + \sum_{i \geq 1} (\varphi_{i+1}^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n - \sum_{i \geq 0} (\varphi_i^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n, \\
= & -(\varphi_1^n)^\top (\pi^0 + \pi^-) D \varphi_1^n + (\varphi_0^n)^\top (\pi^0 + \pi^-) D \varphi_0^n + \left(\frac{1}{2} + \frac{1}{2} \right) \sum_{i \geq 0} (\varphi_{i+1}^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n \\
& \quad - \sum_{i \geq 0} (\varphi_i^n)^\top (\pi^0 + \pi^-) D \varphi_{i+1}^n.
\end{aligned}$$

Arguing the standard indices translation and reorganizing the terms, we deduce

$$\begin{aligned}
& -(\varphi_0^n)^\top(\pi^0 + \pi^-)D(\varphi_1^n - \varphi_0^n) + \sum_{i \geq 1} (\varphi_{i+1}^n - \varphi_i^n)^\top(\pi^0 + \pi^-)D\varphi_{i+1}^n \\
&= -(\varphi_1^n)^\top(\pi^0 + \pi^-)D\varphi_1^n + \frac{1}{2}(\varphi_0^n)^\top(\pi^0 + \pi^-)D\varphi_0^n + \frac{1}{2} \sum_{i \geq 0} (\varphi_{i+1}^n)^\top(\pi^0 + \pi^-)D\varphi_{i+1}^n \\
&\quad + \frac{1}{2} \sum_{i \geq 0} (\varphi_i^n)^\top(\pi^0 + \pi^-)D\varphi_i^n - \sum_{i \geq 0} (\varphi_i^n)^\top(\pi^0 + \pi^-)D\varphi_{i+1}^n, \\
&= -(\varphi_1^n)^\top(\pi^0 + \pi^-)D\varphi_1^n + \frac{1}{2}(\varphi_0^n)^\top(\pi^0 + \pi^-)D\varphi_0^n \\
&\quad + \frac{1}{2} \sum_{i \geq 0} (\varphi_{i+1}^n - \varphi_i^n)^\top(\pi^0 + \pi^-)D(\varphi_{i+1}^n - \varphi_i^n),
\end{aligned}$$

that concludes the proof. □

Acknowledgements

G.B. is supported by the BOURGEONS project, grant ANR-23-CE40- 0014-01 of the French National Research Agency (ANR). We thank Benjamin Boutin for helpful discussion.

References

- [1] Rémi Abgrall, Jan Nordström, Philipp Öffner, and Svetlana Tokareva. Analysis of the SBP-SAT stabilization for finite element methods part I: Linear problems. *Journal of Scientific Computing*, 85:43, 11 2020.
- [2] Rémi Abgrall, Jan Nordström, Philipp Öffner, and Svetlana Tokareva. Analysis of the SBP-SAT stabilization for finite element methods part II: Entropy stability. *Communications on Applied Mathematics and Computation*, 5, 06 2020.
- [3] Geoffrey Beck, Ewan Contentin, and Ludovic Martaud. Freely floating cylinder on a 3d fluid governed by the boussinesq equations. In *16th International conference on mathematical and numerical aspects of wave propagation, WAVES 2024*, 2024.
- [4] Geoffrey Beck and David Lannes. Freely Floating Objects on a Fluid Governed by the Boussinesq Equations. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 39(3):<https://ems.press/journals/aihpc/articles/5300753>, March 2022.
- [5] Geoffrey Beck and David Lannes. Freely floating objects on a fluid governed by the Boussinesq equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, (39):575–646, 2022.
- [6] Geoffrey Beck, David Lannes, and Lisl Weynans. Wave power farm made of many rigid floating structures in Boussinesq regime. In *WAVES 2022 - 15th International Conference on Mathematical and Numerical Aspects of Wave Propagation*, Palaiseau, France, July 2022.
- [7] Geoffrey Beck, David Lannes, and Lisl Weynans. A numerical method for wave-structure interactions in the Boussinesq regime. working paper or preprint, Jul 2023.
- [8] Sylvie Benzoni-Gavage and Denis Serre. *Multi-dimensional hyperbolic partial differential equations: First-order systems and applications*. Oxford University Press, 11 2006.

- [9] Christophe Berthon, Manuel J. Castro Diaz, Arnaud Duran, Tomas Morales De Luna, and Khaled Saleh. Artificial viscosity to get both robustness and discrete entropy inequalities. *Journal of Scientific Computing*, 97(65), 2023.
- [10] Benjamin Boutin. *Finite difference methods for hyperbolic problems with boundaries: stability and multiscale analysis*. Accreditation to supervise research, Université de Rennes, July 2023.
- [11] Benjamin Boutin and Jean-François Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems: numerical boundary layers. *Numerical Mathematics: Theory, Methods and Applications*, 10(3):489–519, 2017.
- [12] Benjamin Boutin, Pierre Le Barbenchon, and Nicolas Seguin. Stability of finite difference schemes for the hyperbolic initial boundary value problem by winding number computations. *Annales de la Faculté des Sciences de Toulouse. Mathématiques.*, 2023.
- [13] Mark Carpenter and Travis Fisher. High-order entropy stable formulations for computational fluid dynamics. In *21st AIAA Computational Fluid Dynamics Conference*, 06 2013.
- [14] Yves Coudière, Jean-Paul Vila, and Philippe Villedieu. Convergence d’un schéma volumes finis explicite en temps pour les systèmes hyperboliques linéaires symétriques en domaines bornés. *Comptes Rendus de l’Académie des Sciences - Series I - Mathematics*, 331(1):95–100, 2000.
- [15] Jean-François Coulombel. Well-posedness of hyperbolic initial boundary value problems. *Journal de mathématiques pures et appliquées*, 84(6):786–818, 2005.
- [16] Jean-François Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems. *SIAM Journal on Numerical Analysis*, 47(4):2844–2871, 2009.
- [17] Jean-François Coulombel. Stability of finite difference schemes for hyperbolic initial boundary value problems ii. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 10(1):37–98, 2011.
- [18] Jean-François Coulombel. Fully discrete hyperbolic initial boundary value problems with nonzero initial data. *Confluentes Mathematici*, 7(2):17–52, 2015.
- [19] François Dubois and Philippe LeFloch. Boundary conditions for nonlinear hyperbolic systems of conservation laws. *Journal of Differential Equations*, 71(1):93–122, 1988.
- [20] Vincent Duchêne. Rigorous justification of the Favrie-Gavrilyuk approximation to the Serre-Green-Naghdi model. *Nonlinearity*, 32(10):3772–3797, 2019.
- [21] Arnaud Duran and Gaël Loïc Richard. Discretization of a new model of dispersive waves with improved dispersive properties and exact conservation of energy. April 2024.
- [22] Marguerite Gisclon and Denis Serre. Conditions aux limites pour un système strictement hyperbolique fournies, par le schéma de Godunov. *ESAIM: Modélisation mathématique et analyse numérique*, 31(3):359–380, 1997.
- [23] Jean-Luc Guermond, Matthias Maier, Bojan Popov, Laura Saavedra, and Ignacio Tomas. First-order greedy invariant-domain preserving approximation for hyperbolic problems: Scalar conservation laws, and p-system. *J. Sci. Comput.*, 100:46, 2023.

- [24] Olivier Guès, Guy Métivier, Mark Williams, and Kevin Zumbrun. Uniform stability estimates for constant-coefficient symmetric hyperbolic boundary value problems. *Communications in Partial Differential Equations*, 32(4):579–590, 2007.
- [25] Amiram Harten, Peter D. Lax, and Bram Van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM review*, 25:35–61, 1983.
- [26] Philippe Helluy, Laurent Navoret, Nhung Pham, and Anaïs Crestetto. Reduced vlasov–maxwell simulations. *Comptes Rendus Mécanique*, 342(10):619–635, 2014. Theoretical and numerical approaches for Vlasov-maxwell equations.
- [27] Tatsuo Iguchi and David Lannes. Hyperbolic free boundary problems and applications to wave-structure interactions. *Indiana University Mathematics Journal*, 2018.
- [28] K. T. Joseph and Philippe LeFloch. Boundary layers in weak solutions of hyperbolic conservation laws. *Archive for Rational Mechanics and Analysis*, 147(1):47–88, June 1999.
- [29] Heinz-Otto Kreiss. Initial boundary value problems for hyperbolic systems. *Communications on Pure and Applied Mathematics*, 23(3):277–298, 1970.
- [30] David Lannes. *The Water Waves Problem: Mathematical Analysis and Asymptotics*, volume 188 of *Mathematical Surveys and Monographs*. 2013.
- [31] David Lannes. Initial boundary value problems for hyperbolic systems, and dispersive perturbations. *Lecture notes of the Bressanone Winter School*, 2018.
- [32] Randall LeVeque. *Numerical methods for conservation laws (2. ed.)*. Lectures in mathematics. Birkhäuser, 1992.
- [33] Andrew J. Majda and S. Osher. Initial-boundary value problems for hyperbolic equations with uniformly characteristic boundary. *Communications on Pure and Applied Mathematics*, 28:607–675, 1975.
- [34] Mohammad Motamed. Generalization of Kreiss theory to hyperbolic problems with boundary-type eigenmodes. *Communications in Mathematical Sciences*, 2019.
- [35] Jan Nordström and Andrew R. Winters. A linear and nonlinear analysis of the shallow water equations and its impact on boundary conditions. *Journal of Computational Physics*, 463:111254, 2022.
- [36] Gaël Richard. An extension of the Boussinesq-type models to weakly compressible flows. *European Journal of Mechanics - B/Fluids*, 89:217–240, September 2021.
- [37] Philip L. Roe. Approximate Riemann solvers, parameter vectors and difference schemes. *J. Comput. Phys.*, 43(2):357–372, 1981.
- [38] V.V. Rusanov. The calculation of the interaction of non-stationary shock waves and obstacles. *USSR Computational Mathematics and Mathematical Physics*, 1(2):304–320, 1962.
- [39] Eitan Tadmor. Numerical viscosity and the entropy condition for conservative difference schemes. *Mathematics of Computation*, 43(168):369–381, 1984.
- [40] Eitan Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. i. *Mathematics of Computation*, 49(179):91–103, 1987.

- [41] Eitan Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12(1):451–512, 2003.
- [42] Eitan Tadmor. Entropy stable schemes. *Handbook of Numerical Methods for Hyperbolic Problems: Basic and Fundamental Issues*, edited by R. Abgrall and C.-W. Shu (North-Holland, Elsevier, Amsterdam, 2017), 17:467–493, 2016.
- [43] Eleuterio F. Toro. *Riemann solvers and numerical methods for fluid dynamics : a practical introduction*. Springer-Verlag, Berlin Heidelberg New York (N.Y.), 2nd ed. edition, cop. 1999.
- [44] Eleuterio F. Toro, Michelle Spruce, and William Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock waves*, 4(1):25–34, 1994.