MATHEMATICS OF COMPUTATION
Volume 87, Number 311, May 2018, Pages 1083–1126
http://dx.doi.org/10.1090/mcom/3253
Article electronically published on September 7, 2017

ENTROPIC SUB-CELL SHOCK CAPTURING SCHEMES VIA JIN-XIN RELAXATION AND GLIMM FRONT SAMPLING FOR SCALAR CONSERVATION LAWS

FRÉDÉRIC COQUEL, SHI JIN, JIAN-GUO LIU, AND LI WANG

ABSTRACT. We introduce a sub-cell shock capturing method for scalar conservation laws built upon the Jin-Xin relaxation framework. Here, sub-cell shock capturing is achieved using the original defect measure correction technique. The proposed method exactly restores entropy shock solutions of the exact Riemann problem and, moreover, it produces monotone and entropy satisfying approximate self-similar solutions. These solutions are then sampled using Glimm's random choice method to advance in time. The resulting scheme combines the simplicity of the Jin-Xin relaxation method with the resolution of the Glimm's scheme to achieve the sharp (no smearing) capturing of discontinuities. The benefit of using defect measure corrections over usual sub-cell shock capturing methods is that the scheme can be easily made entropy satisfying with respect to infinitely many entropy pairs. Consequently, under a classical CFL condition, the method is proved to converge to the unique entropy weak solution of the Cauchy problem for general non-linear flux functions. Numerical results show that the proposed method indeed captures shocks—including interacting shocks—sharply without any smearing.

1. INTRODUCTION

Modern high resolution shock capturing methods for non-linear hyperbolic systems of conservation laws contain two ingredients: building blocks (Godunov type upwind schemes based on exact or approximate Riemann solvers, Lax-Friedrichs type central schemes, kinetic schemes, etc.) [11], [16] and reconstructions that hybridize higher order interpolations in smooth part of the solution and first order methods around discontinuities—shocks and contact discontinuities—(total-variation-diminishing (TVD), essentially-non-oscillatory (ENO) or weighted essentially-non-oscillatory (WENO), discontinuous Galerkin, etc.) [26], [27]. These methods have been very successfully applied to capture shocks in gas dynamics, magnetohydrodynamics, reacting flows and many other problems (see [8] and the references therein). Analyses of these methods, on the other hand, are much less developed and are mostly available only for scalar problems.

Received by the editor June 1, 2015, and, in revised form, November 2, 2016.

2010 *Mathematics Subject Classification.* Primary 35L65, 35L67, 65M12, 76M12, 76M45.

Key words and phrases. Conservation laws, relaxation, defect measure, Riemann problem, shock capturing methods, cell entropy condition, Glimm's random choice.

The second author was partly supported by NSF grants DMS-1522184 and DMS-1107291 (RNMS:KI-Net), and NSFC grant No. 91330203. The research of the third author was partially supported by KI-Net NSF RNMS grant No. 1107444 and NSF grant DMS 0514826. The fourth author was partially supported by NSF grant DMS-1620135. Support for this research was also provided by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation.

Since a first order numerical method necessarily smears shocks, higher order interpolations basically introduce anti-diffusion to counter the smearing of the first order method. The difficulty here is to choose the hybridization in an intelligent way such that the method offers high resolutions near the shock—smears the shock with the fewest number of mesh points—yet does not introduce numerical oscillations and captures the entropic solution of the underlying non-linear hyperbolic systems. One of the earliest modern high resolution schemes aimed at achieving these goals was the Flux-Corrected-Transport (FCT) method [4]. Van Leer introduced the idea of slope limiter [29] to switch between a first order method near the shocks or contact discontinuities and a second order method elsewhere, and then Harten introduced the notion of TVD [12], a generic theory to guide the design of non-oscillatory high resolution shock capturing methods. In all aforementioned high resolution shock *capturing* methods, due to the use of first order methods near discontinuities, the shocks and contact discontinuities are smeared out across a few grid points. Such smearing is not an issue for most inviscid flow calculations. However, there are many problems where the smearing due to numerical viscosities can cause significant pitfalls which lead to polluted or even unphysical numerical solutions. For examples, in multiphase flows, the smeared numerical solutions across the interfaces between the two phases correspond to unphysical phases [20]; in phase transition problems, such as van der Waals flows [28], smeared solutions enter the elliptic regions which are unstable [20]; in the computation of stiff reacting flows [9], the artificially smeared temperature profiles incorrectly trigger chemical reactions which lead to unphysical detonations that propagate with incorrect speeds (see [3] for a correction based on a random projection method). Numerical viscosities are also blamed for numerical oscillations behind slowly moving shocks [17], artificial wall heating [22], and the carbuncle phenomena [24]. Indeed, it contributes to numerical instability in Lax-Friedrichs and Godunov schemes for non-linear hyperbolic systems [1, 2].

This paper aims at developing a one-dimensional shock capturing method that captures shock sharply—*without numerical smearing*—and establishing an entropic convergence theory of this method for scalar conservation laws. The method combines the Jin-Xin relaxation approximation [18] with Dirac measure and Glimm sampling [10]. Thanks to the linear convection of the Jin-Xin relaxation, the Riemann invariants are linear which can be easily inverted and the entropy property satisfied by the scalar conservation laws can be lifted to the relaxation system. We design a specific Dirac measure which allows us to obtain the total-variation-diminishing property and cell entropy condition for both square [23] and Kružkov entropies [19]. The Glimm sampling gives a sharp shock. We refer to [14, 15] for a related sampling strategy based on Roe's approximate solvers and to [5] where a Suliciu solver for the p-system is advocated. In [6], mixed hyperbolic-elliptic Euler equations are solved within the framework of a Sulicu method but using a deterministic front tracking technique, which can also be replaced by a Glimm front sampling. Here, we provide a theoretical foundation for this approach, namely the method indeed converges to the entropic solution of the scalar conservation law for general non-linear fluxes.

Numerically we only use Dirac measure and Glimm sampling near the shock. Elsewhere standard high resolution mechanisms, such as higher order TVD or ENO/WENO reconstruction, can still be used to offer better numerical accuracy.

There were other efforts focused on obtaining sharp shocks numerically. One is the front tracking method which relies on solving Riemann problems exactly. Within the framework of shock capturing methods, which is the approach in this paper, Harten [13] introduced the sub-cell method, which creates an intermediate state based on the conservation property. However, conservation itself only guarantees the capturing of a weak solution according to the Lax-Wendroff theorem. It does not prevent the formation of entropy violating shocks. Our approach always produces entropic shocks.

As with other sub-cell methods, our approach also encounters major challenges when extended to non-linear systems and higher dimensions. This will be a subject of future research.

2. Relaxation defect measures and their numerical application

In this section, we first extend a Dirac measure correction to the Jin-Xin relaxation framework, and then solve the corresponding Riemann problem.

Consider the Cauchy problem for a non-linear scalar conservation law

$$\begin{cases} \partial_t u + \partial_x f(u) = 0,\ t > 0,\ x \in \mathbb{R}, \\ u(t,0) = u_0(x), \end{cases} \tag{2.1}$$

supplemented with the following entropy selection principle:

$$\partial_t \mathcal{U}(u) + \partial_x \mathcal{F}(u) \leq 0. \tag{2.2}$$

Here we assume a smooth flux function $f \in \mathcal{C}^2(\mathbb{R})$, and initial data u_0 is chosen in $L^\infty(\mathbb{R}) \cap \mathrm{BV}(\mathbb{R})$, where BV stands for the space of functions with bounded variation. Inequality (2.2) has to be satisfied in the sense of the distributions for all smooth convex functions $\mathcal{U}(u)$ with $\mathcal{F}'(u) = \mathcal{U}'(u) f'(u)$. It is well-known (see [25] for instance) that the Cauchy problem (2.1)–(2.2) admits a unique entropy weak solution, the so-called Kružkov solution. In [18], Jin and Xin proposed to approximate this solution by that of the following relaxation system,

$$\begin{cases} \partial_t u^\epsilon + \partial_x v^\epsilon = 0, & \text{(2.3a)} \\ \partial_t v^\epsilon + a^2 \partial_x u^\epsilon = -\dfrac{1}{\epsilon}(v^\epsilon - f(u^\epsilon)), & \text{(2.3b)} \end{cases}$$

with well-prepared initial data

$$u(0,x) = u_0(x),\ v(0,x) = v_0(x) = f(u_0(x)). \tag{2.4}$$

Here $\epsilon > 0$ denotes a small relaxation time. For any given fixed $\epsilon > 0$, existence and uniqueness of a solution (u^ϵ, v^ϵ) can be established (see for instance [7], [21]). Furthermore, under the sub-characteristic condition

$$\sup |f'(u)| < a, \tag{2.5}$$

the sequence $\{u^\epsilon, v^\epsilon\}_{\epsilon>0}$ converges strongly as $\epsilon \to 0^+$ in $\mathcal{C}((0,\infty), L^1_{loc}(\mathbb{R}))$ to $(u, f(u))$ with u being the Kružkov solution of (2.1) (see [7], [21] for a precise statement). In particular, this result applies to any initial data u_0 of the form

$$u_0(x) = u_L + (u_R - u_L) H(x), \quad x \in \mathbb{R}, \tag{2.6}$$

where H denotes the Heaviside function, and the constant states u_L and u_R satisfy

$$-\sigma(u_L, u_R)(u_R - u_L) + f(u_R) - f(u_L) = 0 \tag{2.7}$$

and

$$-\sigma(u_L, u_R)\,(\mathcal{U}(u_R) - \mathcal{U}(u_L)) + \mathcal{F}(u_R) - \mathcal{F}(u_L) \le 0 \tag{2.8}$$

for all entropy pairs $(\mathcal{U}, \mathcal{F})$. Here $\sigma(u_L, u_R)$ is the shock speed. This initial data defines a Riemann problem for (2.1) that gives rise to an entropy shock solution

$$u(t,x) = u_L + (u_R - u_L)H(x - \sigma(u_L, u_R)t), \quad t > 0, x \in \mathbb{R}. \tag{2.9}$$

Under the stability condition (2.5), the well-prepared initial data (2.4) built from u_0 in (2.6) leads to a family of solutions $\{(u^\epsilon, v^\epsilon)\}_{\epsilon>0}$ which converges as ϵ goes to zero to $(u, v \equiv f(u))$ where u is given by (2.9). It can be easily shown that the following limit holds in the sense of the distributions

(2.10)
$$\begin{aligned}\lim_{\epsilon\to 0} \frac{1}{\epsilon}(f(u^\epsilon) - v^\epsilon) &= \Big\{ -\sigma(u_L, u_R)(f(u_R) - f(u_L)) + a^2(u_R - u_L)\Big\}\delta_{x-\sigma(u_L,u_R)t} \\ &= (a^2 - \sigma^2(u_L, u_R))(u_R - u_L)\delta_{x-\sigma(u_L,u_R)t}.\end{aligned}$$

Hence the limit pair (u, v) solves again in the sense of the distributions the following system involving a measure source term:

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + a^2\partial_x u = (a^2 - \sigma^2(u_L, u_R))(u_R - u_L)\delta_{x-\sigma(u_L,u_R)t}, \end{cases} \tag{2.11}$$

with initial data

(2.12)
$$u_0(x) = u_L + (u_R - u_L)H(x), \quad v_0(x) := f(u_0(x)) = f(u_L) + (f(u_R) - f(u_L))H(x).$$

In the sequel, the measure source term in (2.11) is referred to as a *relaxation defect measure*.

At this level, it is crucial to observe that although the Cauchy problem (2.3) with the Riemann data (2.12) does not admit a self-similar solution (u^ϵ, v^ϵ) for any given fixed $\epsilon > 0$, the limit PDE model (2.11) does by contrast admit the self-similar solution

$$\begin{aligned} u(t,x) &= u_L + (u_R - u_L)H(x - \sigma(u_L, u_R)t), \quad t > 0,\ x \in \mathbb{R}, \\ v(t,x) &= f(u_L) + (f(u_R) - f(u_L))H(x - \sigma(u_L, u_R)t), \end{aligned} \tag{2.13}$$

where the u-component is nothing but the entropy satisfying shock solution (2.9) of (2.1) and(2.6).

With this in mind, let us briefly revisit the widely used relaxation model (2.3) to the numerical approximation of the Kružkov solution of (2.1). An operator splitting strategy is generally used to circumvent the lack of self-similar solutions. Covering $\mathbb{R}_t^+$ by a collection of small time steps, one thus solves in each time step first the homogeneous Cauchy problem

$$\begin{cases} \partial_t u^\epsilon + \partial_x v^\epsilon = 0, & \text{(2.14a)} \\ \partial_t v^\epsilon + a^2\partial_x u^\epsilon = 0, & \text{(2.14b)} \end{cases}$$

with appropriate initial data, and then the following singular ODE problem

$$\partial_t u^\epsilon = 0, \quad \partial_t v^\epsilon = -\frac{1}{\epsilon}(v^\epsilon - f(u^\epsilon)), \quad \text{in the limit } \epsilon \to 0+, \tag{2.15}$$

again with appropriate data. Here, the first step allows for self-similar solutions, which are made of a single intermediate state $(u^\star, v^\star)$ separated by two discontinuities propagating with speed $-a$ and $+a$, respectively. This step, however, yields a poor resolution of the shock solutions to the original conservation law (2.1). In fact, under the mandatory stability condition (2.5), it is seen [16] that the intermediate value $u^\star$ coincides with the space averaging of the exact self-similar solution (2.9) whose fan is bordered by the two waves $-a$ and $+a$. In other words, the first step of convection will necessarily smear the shock, since the characterstics of (2.14) are not the same as the original equation. In the second step, one can resort to the singular source term $(f(u^\epsilon) - v^\epsilon)/\epsilon$ in the limit $\epsilon \to 0^+$ to adjust to the correct shock speed. Formally speaking, for general well-prepared initial data (2.4), the limit under consideration can be split into two contributions. The first singular part is a Radon measure $\mathcal{M}_{t,x}$, which is the sum of all the relaxation defect measures concentrated on the shocks in the limit solution $u(t,x)$. The second smooth contribution comes from the smooth part of the Kružkov solution and reads $\partial_t f(u) + a^2\partial_x u$. Motivated by this natural decomposition, we propose a new splitting procedure involving in the first step the singular first part $\mathcal{M}_{t,x}$, while the second step is devoted to handle the smooth second part. The first step then consists in solving

$$\begin{cases} \partial_t u^\epsilon + \partial_x v^\epsilon = 0, & \text{(2.16a)} \\ \partial_t v^\epsilon + a^2\partial_x u^\epsilon = \mathcal{M}_{t,x} & \text{(2.16b)} \end{cases}$$

and then

$$\partial_t u^\epsilon = 0, \quad \partial_t v^\epsilon = -\frac{1}{\epsilon}(v^\epsilon - f(u^\epsilon)), \quad \text{in the limit } \epsilon \to 0+, \tag{2.17}$$

with appropriate initial data. Note that the second step cannot develop relaxation defect measures, thus (2.16), (2.17) provide a consistent splitting of the PDE model (2.3) in the limit $\epsilon \to 0^+$. Here, the Cauchy problem (2.16) can be solved by a succession of non-interacting Riemann problems of the form (2.11) and (2.12), once the Radon measure $\mathcal{M}_{t,x}$ is conveniently approximated.

We now describe the main building principle for relevant approximations of $\mathcal{M}_{t,x}$. Consider first the Riemann problems in the generic form

$$\begin{cases} \partial_t u + \partial_x v = 0, & \text{(2.18a)} \\ \partial_t v + a^2\partial_x u = m(u_L, u_R)\delta_{x-\sigma(u_L,u_R)t}, & \text{(2.18b)} \end{cases}$$

with well-prepared initial data

$$\begin{aligned} u(0,x) = u_0(x) &= \begin{cases} u_L, & x < 0, \\ u_R, & x > 0, \end{cases} \\ v(0,x) = v_0(x) &= \begin{cases} f(u_L), & x < 0, \\ f(u_R), & x > 0. \end{cases} \end{aligned} \tag{2.19}$$

In (2.18), $m(u_L, u_R)$ refers to the mass of a Dirac measure concentrated at $x = \sigma(u_L, u_R)t$ where $\sigma(u_L, u_R)$ plays the role of a velocity (2.7) and satisfies

$$|\sigma(u_L, u_R)| < a. \tag{2.20}$$

Both the mass m and velocity σ are to be defined depending on the states u_L, u_R to meet suitable properties for the solution of the Cauchy problem (2.18)–(2.19). But whatever the precise definitions are, the solution we want is clearly self-similar. To condense the notations, $\mathbb{U} = (u, v)^T \in \mathbb{R}^2$ refers to the unknown in (2.18). The case of an identically zero mass $m(u_L, u_R) = 0$ boils down to a Riemann solution for the 2×2 homogeneous linear system (2.14), whose solution consists of three constant states $\mathbb{U}_L$, $\mathbb{U}^\star$ and $\mathbb{U}_R$ separated by two waves propagating with speed $-a$ and $+a$, respectively. For a non-zero mass, easy considerations on the weak form of the PDEs (2.18) reveal the existence of an intermediate discontinuity propagating with speed $\sigma(u_L, u_R)$ (see indeed the condition (2.20)) which separates, in the wave span, two inner states denoted $\mathbb{U}_L^\star$ and $\mathbb{U}_R^\star$. Across the intermediate discontinuity, these two states have to satisfy the following jump conditions:

$$\begin{aligned} &-\sigma(u_L, u_R)(u_R^\star - u_L^\star) + (v_R^\star - v_L^\star) = 0, \\ &-\sigma(u_L, u_R)(v_R^\star - v_L^\star) + a^2(u_R^\star - u_L^\star) = m(u_L, u_R). \end{aligned} \tag{2.21}$$

We then define the mass $m(u_L, u_R)$ and the velocity $\sigma(u_L, u_R)$ in order to preserve some of the essential properties of the exact solution of the Riemann problem

$$\begin{cases} \partial_t u + \partial_x f(u) = 0, \\ u(0, x) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0. \end{cases} \end{cases} \tag{2.22}$$

Properties to be preserved include the *monotonicity* in the self-similar variable $\xi = x/t$, *consistency with the entropy inequalities* (2.2) and *exactness* regarding discontinuous solutions of (2.22).

For pairs of states (u_L, u_R) satisfying (2.7)–(2.8), the exactness amounts to define the mass $m(u_L, u_R)$ and velocity $\sigma(u_L, u_R)$ in the Cauchy problem (2.18)–(2.19) so that its self-similar solution $\mathbb{U}(\xi, u_L, u_R)$ reduces componentwise to (2.13).

In fact, the relaxation defect measure proposed in (2.11) guarantees such a property, as stated in the following lemma.

Lemma 2.1. *Given any pair of states* (u_L, u_R) *verifying the Rankine-Hugoniot condition* (2.7) *and the entropy inequalities* (2.8), *define the velocity*

$$\begin{aligned} &\sigma(u_L, u_R) = \frac{f(u_L) - f(u_R)}{u_L - u_R}, \quad u_L \neq u_R; \\ &\sigma(u_L, u_R) = f'(u_L) = f'(u_R), \quad \textit{otherwise}, \end{aligned} \tag{2.23}$$

and the mass

$$m_s(u_L, u_R) = \left(a^2 - \sigma^2(u_L, u_R)\right)(u_R - u_L). \tag{2.24}$$

Then the solution of the Riemann problem (2.18)–(2.19) $\mathbb{U}(\xi, u_L, u_R)$ *is given by the self-similar function* (2.13).

Proof. One has to prove that the self-similar function (2.13) is a solution of the Riemann problem (2.18)–(2.19) with intermediate states $\mathbb{U}_L^\star = \mathbb{U}_L$ and $\mathbb{U}_R^\star = \mathbb{U}_R$ as long as the velocity and mass are prescribed according to (2.23) and (2.24). It amounts to check that the jump conditions (2.21) across the intermediate wave are satisfied with $\mathbb{U}_L^\star = \mathbb{U}_L$ and $\mathbb{U}_R^\star = \mathbb{U}_R$. But, clearly

$$-\sigma(u_L, u_R)(u_R^\star - u_L^\star) + (v_R^\star - v_L^\star) = -\sigma(u_L, u_R)(u_R - u_L) + f(u_R) - f(u_L) = 0, \tag{2.25}$$

by the definition of $\sigma(u_L, u_R)$ while the mass $m_s(u_L, u_R)$ prescribed in (2.24) satisfies the identity:

$$\begin{aligned} &-\sigma(u_L, u_R)(v_R^\star - v_L^\star) + a^2(u_R^\star - u_L^\star) \\ &\qquad = -\sigma^2(u_L, u_R)(u_R - u_L) + a^2(u_R - u_L) := m_s(u_L, u_R). \end{aligned} \tag{2.26}$$

This concludes the proof. □

Notice that the definition of mass $m_s(u_L, u_R)$ (2.24) only takes into account the jump condition (2.7), but not the entropy condition(s) (2.8). Since it is always possible to define the velocity $\sigma(u_L, u_R)$ satisfying (2.7) for any given pair of states

(u_L, u_R), the choice of $m_s(u_L, u_R)$ in (2.24) would systematically result in a solution given by (2.13), which does not necessarily satisfy the entropy condition. In other words, we would merely end up with a Roe scheme which is known to be entropy violating in the approximation of the solutions of (2.22) [16].

We thus propose to modulate the definition of the mass in (2.24) by looking for a monitoring factor $\theta(u_L, u_R)$, namely a suitable real-valued mapping $\theta : (u_L, u_R) \in \mathbb{R}^2 \to \theta(u_L, u_R) \in \mathbb{R}$, to define

$$m(u_L, u_R) = \theta(u_L, u_R) \left(a^2 - \sigma^2(u_L, u_R)\right) (u_R - u_L), \tag{2.27}$$

with $\sigma(u_L, u_R)$ given by (2.23). Clearly, $\theta(u_L, u_R)$ acts as an anti-diffusive parameter, allowing for a continuous shift from the Lax Friedrichs scheme when $\theta(u_L, u_R) = 0$ to the Roe scheme for $\theta(u_L, u_R) = 1$.

2.1. **Design principle of approximate defect measures.** We now briefly explain the design principle for relevant anti-diffusive laws $\theta(u_L, u_R)$. First, we consider the case of a strictly convex flux function $f(u)$. An immediate choice for the anti-diffusive law $\theta(u_L, u_R)$ would be

$$\theta(u_L, u_R) = \begin{cases} 1, & u_L > u_R, \\ 0, & \text{otherwise}, \end{cases} \tag{2.28}$$

since the situation $u_L > u_R$ yields an entropy satisfying shock solution, while the converse gives rise to a rarefaction. Actually we will prove that more anti-diffuse choices for $\theta(u_L, u_R)$ can be performed while still allowing for convergence to the Kružkov solution. In particular, we will prove that $\theta(u_L, u_R)$ can be set close to 1 (at the order $\mathcal{O}(\Delta x)$ with $\Delta x > 0$ the space step) in the smooth part of the approximate solution. Hence rarefaction waves in the discrete solution can be handled with θ asymptotically close to 1 (not 0) as advocated in (2.28).

The derivation of relevant anti-diffusive laws θ essentially relies on a consistency requirement with the entropy inequalities (2.2). In the case of a genuinely nonlinear flux $f(u)$, it is known after Panov [23] that a single strictly convex entropy pair suffices to select the Kružkov solution of (2.1). θ-laws are derived accordingly on the ground of a single entropy pair. The situation of a general non-linear flux function is more involved. First, the obvious choice (2.28) no longer applies. Second, infinitely many entropy pairs are required to single out the Kružkov solution. We are thus led to design θ-laws accordingly by considering infinitely many entropy pairs.

Our consistency condition with the entropy inequalities (2.2) is built from the relaxation entropy pairs associated with the Jin-Xin's model (2.18). As established in [7] (see also [21]), any given smooth convex entropy pair $(\mathcal{U}, \mathcal{F})$ for (2.1) can be suitably lifted to a relaxation entropy pair for (2.18), which we denote (Φ, Ψ) in the sequel. Under the sub-characteristic condition (2.5), the relaxation mechanism in (2.18) can be shown to be dissipative with respect to any of those relaxation entropy pairs. More precisely, given (2.5), an invariant domain exists for the solutions of (2.18), within which convexity and dissipative properties for any pair (Φ, Ψ) can be proved. As these two crucial properties are generically lost outside of the invariant domain, it is of central importance to keep such a domain invariant for the solution of the Riemann problem (2.18). This requirement will be fairly easy to achieve from the choice (2.27), allowing us in turn to enforce consistency with the entropy inequalities (2.2).

2.2. **Organization of the paper.** In section 2.3, the Riemann problem (2.18)–(2.19) with a defect measure correction (2.27) is solved for a general pair of states (u_L, u_R). A central property (see Corollary 2.3) due to the choice (2.27) is then revealed in the characteristic variables $(v - av\ , u + av)$, which allows us to prove in section 3 the existence of entropy invariant domain for self-similar solutions of (2.18) provided that the anti-diffusive law $\theta(u_L, u_R)$ takes values in $[0, 1]$. Equivalence of such invariance property with a monotonicity property for u-component of the solution (2.18) is then established. As a consequence, uniform sup-norm and BV estimates are inferred, allowing us to prove the convergence of the Jin-Xin relaxation solver with defect measure correction to a weak solution of (2.1). To enforce the entropy condition with the expected Kružkov solution, in subsection 3.2, we require that the discontinuity induced by the approximate defect measure in (2.18) is entropy satisfying with respect to relaxation entropy pairs (Φ, Ψ). This requirement further confines the admissible graph of relevant anti-dissipative law $\theta : (u_L, u_R) \in \mathbb{R}^2 \to [0, \Theta(u_L, u_R)]$, where the positive real number $\Theta(u_L, u_R)$ denotes some optimal upper-bound. $\Theta(u_L, u_R)$ is in general strictly less than 1 for arbitrary pairs (u_L, u_R) and is equal to 1 for the pairs satisfying (2.8). In other words, *exact* capturing of entropy shock solutions is thus assured. In subsection 3.3, we perform the analysis for a strictly convex flux on the ground of a single entropy inequality. This analysis is extended in subsection 3.4 to general flux functions, involving the whole family of Kružkov entropy pairs. In both settings, the optimal upper-bound $\Theta(u_L, u_R)$ is given explicitly. In section 4, two numerical methods for approximating the Kružkov solution of (2.1)–(2.2) are introduced. The convergence of the corresponding families of approximate solutions is established in section 5. At last, we give some numerical results to assess the critical importance of designing optimal anti-diffusive law $\Theta(u_L, u_R)$ according to infinitely many entropy pairs in the frame of a flux function without genuine non-linearities. Numerical results show that the proposed method indeed captures shocks—including interacting shocks—sharply without any smearing.

2.3. **Solution to the relaxation Riemann problem with defect measure correction.** Consider a pair of real numbers (u_L, u_R) and a constant positive velocity a prescribed under the sub-characteristic condition

$$\sup_{u \in \lfloor u_L, u_R \rceil} |f'(u)| < a, \tag{2.29}$$

where $\lfloor a, b \rceil$ denotes in the sequel the interval $[\min(a, b), \max(a, b)]$ for any given pair of real numbers (a, b). We first give the precise form of the self-similar solution of the Riemann problem (2.18)–(2.19).

Proposition 2.2. *Define the velocity* $\sigma(u_L, u_R)$ *according to* (2.23) *and consider a mass* $m(u_L, u_R)$ *under the form* (2.27) *for some given mapping* $\theta : (u_L, u_R) \in \mathbb{R}^2 \to \theta(u_L, u_R) \in \mathbb{R}$. *Then the solution* $\mathbb{U}(.; u_L, u_R)$ *of the Riemann problem* (2.18)–(2.19) *consists of four constant states* $\mathbb{U}_L$, $\mathbb{U}_L^\star(\theta; u_L, u_R)$, $\mathbb{U}_R^\star(\theta; u_L, u_R)$ *and* $\mathbb{U}_R$ *separated by three discontinuities propagating with speed* $-a$, $\sigma(u_L, u_R)$ *and* $+a$, *respectively. Define*

$$u^\star = \frac{1}{2}(u_L + u_R) - \frac{1}{2a}\big(f(u_R) - f(u_L)\big), \quad v^\star = \frac{1}{2}\big(f(u_R) + f(u_L)\big) - \frac{a}{2}(u_R - u_L). \tag{2.30}$$

Then the intermediate state $\mathbb{U}_L^\star(\theta; u_L, u_R)$ *reads componentwisely as*

$$\begin{aligned} u_L^\star(\theta; u_L, u_R) &= u^\star - \frac{1}{2a}\theta(u_L, u_R)(a - \sigma(u_L, u_R))(u_R - u_L), \\ v_L^\star(\theta; u_L, u_R) &= v^\star + \frac{1}{2}\theta(u_L, u_R)(a - \sigma(u_L, u_R))(u_R - u_L); \end{aligned} \tag{2.31}$$

while $\mathbb{U}_R^\star(\theta; u_L, u_R)$ *is given by*

$$\begin{aligned} u_R^\star(\theta; u_L, u_R) &= u^\star + \frac{1}{2a}\theta(u_L, u_R)(a + \sigma(u_L, u_R))(u_R - u_L), \\ v_R^\star(\theta; u_L, u_R) &= v^\star + \frac{1}{2}\theta(u_L, u_R)(a + \sigma(u_L, u_R))(u_R - u_L). \end{aligned} \tag{2.32}$$

Proof. The two jump conditions (2.21) at the intermediate discontinuity are supplemented with Rankine-Hugoniot relations for the waves propagating with speed $-a$ and $+a$, respectively,

$$a(u_L^\star - u_L) + (v_L^\star - f(u_L)) = 0, \quad -a(u_R - u_R^\star) + (f(u_R) - v_R^\star) = 0. \tag{2.33}$$

The resulting 4×4 non-linear system then has a unique solution for any given mass m provided that $|\sigma(u_L, u_R)| \neq a$, which holds due to the sub-characteristic condition (2.20) with the choice (2.23). With little abuse in the notation, the components of the intermediate states read

$$\begin{aligned} u_L^\star(m) &= u^\star - \frac{m}{2a(a+\sigma)}, \quad v_L^\star(m) = v^\star + \frac{m}{2(a+\sigma)}, \\ u_R^\star(m) &= u^\star + \frac{m}{2a(a-\sigma)}, \quad v_R^\star(m) = v^\star + \frac{m}{2(a-\sigma)}, \end{aligned} \tag{2.34}$$

with $u^\star$ and $v^\star$ given in (2.30). The required expressions (2.31)–(2.32) readily follow after plugging in the particular form (2.27) for the mass under consideration. □

Notice that the state $\mathbb{U}^\star \equiv (u^\star, v^\star)$ defined in (2.30) is nothing but the intermediate state involved in the classical solution for the homogeneous Riemann problem (2.18)–(2.19), i.e., with $m(u_L, u_R) = 0$. In the next few sections, we will show that the monitoring weight functions $\theta(u_L, u_R)$ naturally keep their values in the interval $[0, 1]$, which makes the intermediate states $\mathbb{U}_L^\star$ and $\mathbb{U}_R^\star$ well behaved.

A central property due to the choice (2.27) of the mass $m(u_L, u_R)$ is revealed in the following corollary when reformulating the two intermediate states using the characteristic variables

$$r^\pm = v \pm au. \tag{2.35}$$

Corollary 2.3. *Under the assumptions of Proposition* 2.2*, we re-express the intermediate states* $\mathbb{U}_L^\star(\theta; u_L, u_R)$ *and* $\mathbb{U}_R^\star(\theta; u_L, u_R)$ *into the characteristic variables*

$$r_L^{\pm\star}(\theta) = v_L^\star(\theta; u_L, u_R) \pm a u_L^\star(\theta; u_L, u_R), \quad r_R^{\pm\star}(\theta) = v_R^\star(\theta; u_L, u_R) \pm a u_R^\star(\theta; u_L, u_R). \tag{2.36}$$

Then $r_L^{-\star}(\theta)$ *and* $r_R^{+\star}(\theta)$ *can be equivalently rewritten as linear combinations in* θ *of* $r_L^\pm = f(u_L) \pm a u_L$ *and* $r_R^\pm = f(u_R) \pm a u_R$ *according to*

$$\begin{aligned} r_L^{-\star}(\theta) &= \theta(u_L, u_R) r_L^- + \big(1 - \theta(u_L, u_R)\big) r_R^-, \\ r_R^{+\star}(\theta) &= \big(1 - \theta(u_L, u_R)\big) r_L^+ + \theta(u_L, u_R) r_R^+, \end{aligned} \tag{2.37}$$

where we have

$$r_R^{-\star}(\theta) = r_R^-, \quad r_L^{+\star}(\theta) = r_L^+. \tag{2.38}$$

The proof is straightforward, so we leave it to the reader. As already claimed, relevant mappings $\theta(u_L, u_R)$ will be shown to keep values in the interval $[0, 1]$. In this respect, the linear combinations in (2.37) are nothing but convex decompositions of the left and right data expressed in terms of the characteristic variables (2.35). Such convex decompositions is crucial in the design of explicit mappings $\theta(u_L, u_R)$ when considering the consistency conditions with the entropy inequalities (2.2) in the forthcoming sections.

3. Design of non-linearly stable mappings $\theta(u_L, u_R)$

This section studies the monitoring mapping $\theta(u_L, u_R)$ in (2.27) for general pairs of states (u_L, u_R) so that the solution of the Riemann problem with defect measure correction (2.18)–(2.19) obeys linear and non-linear stability properties.

3.1. **Monotonicity preservation.** The main result of this subsection follows.

Proposition 3.1. *For a general pair of states* (u_L, u_R)*, define the velocity* $\sigma(u_L, u_R)$ *according to* (2.23) *and consider a mass* $m(u_L, u_R)$ *under the form* (2.27)*. Then under the sub-characteristic condition* (2.29)*, the u-component of the Riemann solution* $\mathbb{U}(.; u_L, u_R)$ *of the problem* (2.18)*–*(2.19) *satisfies the monotonicity preserving properties*

$$TV(u(\cdot\ ; u_L, u_R)) = |u_R - u_L| \tag{3.1}$$

if and only if

$$0 \leq \theta(u_L, u_R) \leq 1. \tag{3.2}$$

As a consequence, we have

$$\min(u_L, u_R) \leq u(\cdot\ ; u_L, u_R) \leq \max(u_L, u_R) \tag{3.3}$$

and, moreover,

$$|v(\cdot\ ; u_L, u_R)| \leq 2a \max(|u_L|, |u_R|), \quad TV\big(v(\cdot\ ; u_L, u_R)\big) \leq a|u_L - u_R|. \tag{3.4}$$

Proof. We will consider the case $u_L < u_R$, the reverse situation follows similar steps. If $\theta(u_L, u_R)$ is defined so that the following ordering is valid,

$$u_L \leq u_L^\star(\theta\ ; u_L, u_R) \leq u_R^\star(\theta\ ; u_L, u_R) \leq u_R, \tag{3.5}$$

then the total variation estimate stated in (3.1) is guaranteed, so is the maximum principle (3.3). Indeed, using the definitions of the intermediate values in (2.30)–(2.32), one gets

$$\begin{aligned}
u_L^\star(\theta) - u_L &= (u^\star - u_L) - \theta(u_L, u_R)\frac{a-\sigma}{2a}(u_R - u_L) \\
&= \big(1 - \theta(u_L, u_R)\big)\frac{a-\sigma}{2a}(u_R - u_L), \\
u_R^\star(\theta) - u_L^\star(\theta) &= \theta(u_L, u_R)(u_R - u_L), \\
u_R - u_R^\star(\theta) &= (u_R - u^\star) - \theta(u_L, u_R)\frac{a+\sigma}{2a}(u_R - u_L) \\
&= \big(1 - \theta(u_L, u_R)\big)\frac{a+\sigma}{2a}(u_R - u_L).
\end{aligned} \tag{3.6}$$

Therefore, under the sub-characteristic condition (2.20) inherited from (2.29), the proposed ordering (3.5) holds if and only if the weight $\theta(u_L, u_R)$ verifies (3.2).

Regarding the sup-norm estimate for the second component $v(.\,;u_L,u_R)$, we notice that

$$\text{sign}(u_R-u_L)v_L^\star(\theta)=\text{sgn}(u_R-u_L)v^\star+\frac{\theta}{2}(a-\sigma)|u_R-u_L|,$$

which is an increasing function of $\theta\in[0,1]$:

$$\begin{aligned}\text{sign}(u_R-u_L)v^\star&\le\text{sign}(u_R-u_L)v_L^\star(\theta)\\&\le\text{sign}(u_R-u_L)\big(v^\star+\frac{1}{2}(a-\sigma)(u_R-u_L)\big),\\&=\text{sign}(u_R-u_L)f(u_L).\end{aligned}\tag{3.7}$$

We then have

(3.8)

$$|v_L^\star(\theta)|\le\max(|f(u_L)|,|v^\star|),\quad\text{with } |v^\star|\le\max(|f(u_L)|,|f(u_R)|)+a/2|u_R-u_L|,$$

and assuming without loss of generality that $f(0)=0$, we infer under the sub-characteristic condition (2.29) the desired estimate $|v_L^\star(\theta)|\le 2a\max(|u_L|,|u_R|)$. The same estimate holds for $|v_R^\star(\theta)|$. At last, the total variation of $v(.;u_L,u_R)$ reads

$$\begin{aligned}\text{TV}\big(v(\cdot;u_L,u_R)\big)&=|v_L^\star(\theta)-v_L|+|v_R^\star(\theta)-v_L^\star(\theta)|+|v_R-v_R^\star(\theta)|\\&=\big((1-\theta)a+\theta|\sigma|\big)|u_R-u_L|\\&\le a|u_R-u_L|,\end{aligned}$$

where we have used the definitions of the intermediate states (2.30)–(2.32) and the sub-characteristic condition (2.20). □

As is well-known, the solution $u(.;u_L,u_R)$ of the Riemann problem (2.22) and (2.2) satisfies the *a priori* estimates (3.1). Therefore, it is natural to require the u-component of $\mathbb{U}(.;u_L,u_R)$ to satisfy the same estimates. This in turn requires $\theta(u_L,u_R)$ to satisfy the condition (3.2). We would like to emphasis that condition (3.2) actually implies a stronger property for $\mathbb{U}(.;u_L,u_R)$. To be more specific, we will adopt a broader viewpoint.

After Chen, Levermore and Liu [7] and Natalini [21], define the following two functions,

$$h_\pm(u)=f(u)\pm au,\quad u\in\lfloor u_L,u_R\rceil,\tag{3.9}$$

and consider the compact intervals $\mathcal{I}_-=h_-(\lfloor u_L,u_R\rceil)$ and $\mathcal{I}_+=h_+(\lfloor u_L,u_R\rceil)$. Under the sub-characteristic condition (2.29), the inverse functions $h_\pm^{-1}:r\in\mathcal{I}_\pm\to h_\pm^{-1}(r)\in\lfloor u_L,u_R\rceil$ are well-defined with the property that h_+^{-1} (respectively, h_-^{-1}) is increasing (respectively, decreasing)

$$\frac{d}{dr}h_+^{-1}(r)=\frac{1}{a+f'(h_+^{-1}(r))}>0,\quad\frac{d}{dr}h_-^{-1}(r)=-\frac{1}{a-f'(h_-^{-1}(r))}<0.\tag{3.10}$$

Equipped with these notations, we built the following compact domain of $\mathbb{R}^2$ from the interval $\lfloor u_L,u_R\rceil$

(3.11)

$$\mathcal{D}(\lfloor u_L,u_R\rceil)\equiv\{\mathbb{U}=(u,v)\in\mathbb{R}^2;\ r_-(\mathbb{U})=v-au\in\mathcal{I}_-\text{ and } r_+(\mathbb{U})=v+au\in\mathcal{I}_+\}.$$

Of critical importance in the sequel, the domain (3.11) can be shown to stay invariant by the Jin-Xin relaxation model under the sub-characteristic condition (2.29) (see [7], [21]). Namely given a well-prepared initial data $\mathbb{U}_0=(u_0,v_0=f(u_0))$

with $u_0(x) \in \lfloor u_L, u_R \rceil$ for a.e. $x \in \mathbb{R}$, then for any given relaxation time $\epsilon > 0$, the unique solution $\mathbb{U}^\epsilon$ of the Cauchy relaxation problem stays in $\mathcal{D}(\lfloor u_L, u_R \rceil)$. In particular, the solution of Riemann problem of the homogeneous system (2.18), i.e., with $m(u_L, u_R) = 0$, satisfies this invariance property. This property turns out to be crucial in the dissipative convex lift of the convex entropy pairs $(\mathcal{U}, \mathcal{F})$ for (2.22) to the relaxation system. Condition (3.2) also guarantees the invariance property for the the corresponding solutions of the Riemann problem of the system with defect measure correction, as shown by the next corollary.

Corollary 3.2. *Given a pair of states* (u_L, u_R)*, assume the sub-characteristic condition* (2.29)*. Then the solution* $\mathbb{U}(., u_L, u_R)$ *of* (2.18)–(2.19) *for a given* $\theta(u_L, u_R)$ *stays in* $\mathcal{D}(\lfloor u_L, u_R \rceil)$ (3.11) *if and only if the monotonicity preserving condition* (3.2) *is satisfied.*

Proof. This statement is a direct consequence of Corollary 2.3. Indeed, under the sub-characteristic condition (2.29), r_L^-, r_R^- (respectively, r_L^+, r_R^+) are nothing but the boundaries of the interval $\mathcal{I}_-$ (respectively, $\mathcal{I}_+$) in view of the monotonicity properties (3.10). Keeping the domain $\mathcal{D}(\lfloor u_L, u_R \rceil)$ invariant is thus equivalent to require that the characteristic variables $r_L^{-\star}(\theta; u_L, u_R)$, $r_R^{-\star}(\theta; u_L, u_R)$ are convex combinations of these two boundaries (respectively, $r_L^{+\star}(\theta; u_L, u_R)$, $r_R^{+\star}(\theta; u_L, u_R)$). According to (2.37), such a property is met if and only if the monotonicity preserving condition (3.2) is met. □

3.2. **Entropy consistency requirements.** In this section, we propose and analyze entropy-like conditions to further restrict the graph of the monotonicity preserving mapping $\theta(u_L, u_R)$. In particular, we require $\theta(u_L, u_R) = 1$ for pairs of states (u_L, u_R) that satisfy entropy inequalities (2.8). As already underlined, the entropy consistency condition we consider concerns a single entropy pair

$$\mathcal{U}(u) = \frac{u^2}{2}, \quad \mathcal{F}(u) = \int_0^u v f'(v) dv \tag{3.12}$$

in the case of a genuinely non-linear flux function $f(u)$; and the Kružkov family of entropy pairs

$$\mathcal{U}_k = |u - k|, \quad \mathcal{F}_k(u) = \operatorname{sign}(u - k)\big(f(u) - f(k)\big), \quad k \in \mathbb{R}, \tag{3.13}$$

in the case of a general non-linear flux. Our entropy consistency requirement relies on the extension of entropy pairs proposed in [7, 21] to the Jin-Xin relaxation system from convex entropy pairs of the scalar conservation law (2.22). Here we briefly revisit their design principle as it is of importance hereafter.

For any given interval of the form $\lfloor u_L, u_R \rceil$, the proposed extension is performed over the compact domain $\mathcal{D}(\lfloor u_L, u_R \rceil)$ defined in (3.11). In [7, 21], suitable properties for the proposed extension actually follow from the invariance property of $\mathcal{D}(\lfloor u_L, u_R \rceil)$ under the sub-characteristic condition (2.29). Such an invariance property indeed guarantees the monotonicity properties (3.10) of the functions $h^\pm$ defined in (3.9) for states $\mathbb{U}$ in $\mathcal{D}(\lfloor u_L, u_R \rceil)$. Let us stress that in the present setting, those properties are equivalently preserved under the monotonicity preserving condition (3.2) as put forward in Corollary 3.2.

Given an entropy pair $(\mathcal{U}, \mathcal{F})$ for the scalar equation (2.22), one seeks an entropy pair (Φ, Ψ) for the Jin-Xin relaxation equations which is well defined over

$\mathcal{D}(\lfloor u_L, u_R \rceil)$ and coincides with $(\mathcal{U}, \mathcal{F})$ at equilibrium, namely,

$$\Phi(u, f(u)) = \mathcal{U}(u), \ \Psi(u, f(u)) = \mathcal{F}(u), \quad \text{for all } u \in \lfloor u_L, u_R \rceil. \tag{3.14}$$

General entropy pairs for the (homogeneous) Jin-Xin relaxation equations read

$$\begin{aligned} \Phi(\mathbb{U}) &= \varphi^+(r^+(\mathbb{U})) + \varphi^-(r^-(\mathbb{U})), \\ \Psi(\mathbb{U}) &= a\big(\varphi^+(r^+(\mathbb{U})) - \varphi^-(r^-(\mathbb{U}))\big), \end{aligned} \tag{3.15}$$

with $r^{\pm}(\mathbb{U}) = v \pm au$ for arbitrary pairs of functions (φ^-, φ^+). The consistency requirement (3.14) is therefore met if and only if

$$\begin{aligned} &\varphi^-(h_-(u)) = \frac{1}{2}\Big(\mathcal{U}(u) - \frac{1}{a}\mathcal{F}(u)\Big), \\ &\varphi^+(h_+(u)) = \frac{1}{2}\Big(\mathcal{U}(u) + \frac{1}{a}\mathcal{F}(u)\Big) \quad \text{for all } u \in \lfloor u_L, u_R \rceil, \end{aligned} \tag{3.16}$$

where $h^{\pm}(u)$ denote the two functions introduced in (3.9). Observe that as a consequence the functions $\varphi^{\pm} : r \in \mathcal{I}_{\pm} \to \varphi^{\pm}(r) \in \mathbb{R}$ under consideration satisfy

$$\frac{d}{dr}\varphi^+(r) = \frac{1}{2a}\mathcal{U}'(h_+^{-1}(r)), \ \ \frac{d}{dr}\varphi^-(r) = -\frac{1}{2a}\mathcal{U}'(h_-^{-1}(r)), \tag{3.17}$$

where again $h_{\pm}^{-1} : r \in \mathcal{I}_{\pm} \to h_{\pm}^{-1}(r) \in \lfloor u_L, u_R \rceil$ are well-defined under the sub-characteristic condition (2.29). Due to the convexity of $\mathcal{U}(u)$, the monotonicity properties (3.10) of $h_{\pm}^{-1}$ then ensures the convexity of $\Phi(\mathbb{U})$ over the domain $\mathcal{D}(\lfloor u_L, u_R \rceil)$. Observe that the definitions (3.16) for $\varphi^{\pm}$ are meaningful in the case of the piecewise smooth Kružkov entropies (3.13).

Equipped with (3.15)–(3.17), one then investigates the dissipative properties of the convex extension (Φ, Ψ) with respect to the relaxation mechanisms involved in the Jin-Xin's model. It can be shown (see again [7,21]) that provided the compact domain $\mathcal{D}(\lfloor u_L, u_R \rceil)$ stays invariant for the relaxation equations, we have

$$\partial_v \Phi(u, v)(f(u) - v) \leq 0 \quad \text{for any given } \mathbb{U} = (u, v) \in \mathcal{D}(\lfloor u_L, u_R \rceil), \tag{3.18}$$

which implies that for all relaxation time $\epsilon > 0$, the solutions $\mathbb{U}^\epsilon$ of the Jin-Xin relaxation model with well-prepared initial data $\mathbb{U}_0$ taking values in $\mathcal{D}(\lfloor u_L, u_R \rceil)$ obey the entropy-like inequality

$$\partial_t \Phi(\mathbb{U}^\epsilon) + \partial_x \Psi(\mathbb{U}^\epsilon) = \frac{1}{\epsilon}\partial_v \Phi(\mathbb{U}^\epsilon)(f(u^\epsilon) - v^\epsilon) \leq 0 \tag{3.19}$$

in the usual weak sense. Recall that $\mathbb{U}^\epsilon$ remains in the invariant region $\mathcal{D}(\lfloor u_L, u_R \rceil)$ in view of Corollary 3.2, we now examine the behavior of the relaxation entropy pair (Φ, Ψ) for the self-similar solution $\mathbb{U}(.; u_L, u_R)$ of the Riemann problem (2.18)–(2.19). Note first that $\mathbb{U}(.; u_L, u_R)$ stays constant except across three discontinuities. Concerning the two waves with speed $-a$ and $+a$, their linear degeneracy ensures [25] that any given additional entropy condition is exactly preserved for weak solutions. Namely, whatever the pair of states (u_L, u_R) are and the definition of the mapping θ under (3.2) is, one has

$$\begin{aligned} a\big(\Phi(\mathbb{U}_L^\star(\theta; u_L, u_R)) - \Phi(\mathbb{U}_L)\big) + \Psi(\mathbb{U}_L^\star(\theta; u_L, u_R)) - \Psi(\mathbb{U}_L) = 0, \\ -a\big(\Phi(\mathbb{U}_R) - \Phi(\mathbb{U}_R^\star(\theta; u_L, u_R))\big) + \Psi(\mathbb{U}_R) - \Psi(\mathbb{U}_R^\star(\theta; u_L, u_R)) = 0. \end{aligned} \tag{3.20}$$

Here $\mathbb{U}_L^\star(\theta; u_L, u_R)$ and $\mathbb{U}_R^\star(\theta; u_L, u_R)$ denote the two intermediate states in (2.31)–(2.32) separated by the discontinuity propagating with speed $\sigma(u_L, u_R)$. At this discontinuity, the defect measure correction comes into play. The inequality (3.19) suggests that $\theta(u_L, u_R)$ should satisfy the entropy-like jump condition

$$\begin{aligned}\mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R) := &-\sigma(u_L, u_R)\big(\Phi(\mathbb{U}_R^\star(\theta; u_L, u_R)) - \Phi(\mathbb{U}_L^\star(\theta; u_L, u_R))\big)\\ &+ \Psi(\mathbb{U}_R^\star(\theta; u_L, u_R)) - \Psi(\mathbb{U}_L^\star(\theta; u_L, u_R))\\ \le{}& 0\end{aligned} \tag{3.21}$$

for any given pair of states (u_L, u_R). These observations motivate the following.

Definition 3.3. Given any convex entropy pair $(\mathcal{U}, \mathcal{F})$ (2.2) for the scalar conservation law (2.22) and its relaxation extension (Φ, Ψ) (3.15)–(3.17), the monotonicity preserving mapping θ in (2.27) is said to be consistent with $(\mathcal{U}, \mathcal{F})$ if the relaxation entropy jump $\mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R)$ defined in (3.21) is non-positive for all pair of states (u_L, u_R).

Notice that choosing $\theta(u_L, u_R) = 1$ for special pairs (u_L, u_R) satisfying (2.8) is allowed by the proposed condition. Indeed, Lemma 2.1 ensures that $\mathbb{U}_L^\star(1; u_L, u_R) = \mathbb{U}_L$ and $\mathbb{U}_R^\star(1; u_L, u_R) = \mathbb{U}_R$. Since $\mathbb{U}_L$, $\mathbb{U}_R$ is well-prepared (2.19), the consistency property (3.14) relating (Φ, Ψ) to $(\mathcal{U}, \mathcal{F})$ readily implies

$$\begin{aligned}\mathcal{E}\{\mathcal{U}\}(1; u_L, u_R) &= -\sigma(u_L, u_R)\big(\Phi(\mathbb{U}_R^\star(1; u_L, u_R)) - \Phi(\mathbb{U}_L^\star(1; u_L, u_R))\big)\\ &\quad + \Psi(\mathbb{U}_R^\star(1; u_L, u_R)) - \Psi(\mathbb{U}_L^\star(1; u_L, u_R))\\ &= -\sigma(u_L, u_R)\big(\Phi(\mathbb{U}_R) - \Phi(\mathbb{U}_L)\big) + \Psi(\mathbb{U}_R) - \Psi(\mathbb{U}_L)\\ &= -\sigma(u_L, u_R)\big(\mathcal{U}(u_R) - \mathcal{U}(u_L)\big) + \mathcal{F}(u_R) - \mathcal{F}(u_L)\\ &\le 0.\end{aligned} \tag{3.22}$$

Hence, the entropy condition (3.21) is automatically satisfied by states satisfying (2.8). For general states (u_L, u_R), Definition 3.3 will be used in connection with the following lemma which states that the minimum in the v-variable of any strictly convex relaxation entropy $\Phi(u, v)$ lies on the equilibrium manifold. It thus restores the equilibrium entropy $\mathcal{U}(u)$.

Lemma 3.4. *Assume the sub-characteristic condition* (2.29)*, one has for any given* $u \in \lfloor u_L, u_R \rceil$ *the following Gibb's principle:*

$$f(u) = \operatorname{argmin}_v \Phi(u, v). \tag{3.23}$$

Proof. Let u be given in $\lfloor u_L, u_R \rceil$, then by convexity of $\mathcal{U}(u)$, solving in v the equation

$$\partial_v \Phi(u, v) = \frac{1}{2a}\Big(\mathcal{U}'(h_+^{-1}(v + au)) - \mathcal{U}'(h_-^{-1}(v - au))\Big) = 0,$$

is equivalent to

$$h_+^{-1}(v + au) - h_-^{-1}(v - au) = 0.$$

Under condition (2.29) and for all $(u, v) \in \mathcal{D}(\lfloor u_L, u_R \rceil)$, the function $G(v) = h_+^{-1}(v+au) - h_-^{-1}(v-au)$ is strictly increasing in v thanks to (3.10), thus the unique solution of $G(v) = 0$ is given by $v = f(u)$ since $h_+^{-1}(f(u)+au) = h_-^{-1}(f(u)-au) = u$. Then the equality $\mathcal{U}(u) = \Phi(u, f(u))$ gives the conclusion. □

3.3. **Entropy consistency for genuinely non-linear flux functions.** The main result of this section follows.

Theorem 3.5. *Consider the entropy pair* $(\mathcal{U}(u), \mathcal{F}(u))$ (2.2) *with* $\mathcal{U}(u) = u^2/2$ *and the associated relaxation entropy pair* (Φ, Ψ) (3.15)–(3.16). *Assume the sub-characteristic condition* (2.29). *Then the monotonicity preserving condition* (3.1) *and the entropy condition* $\mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R) \le 0$ (3.21) *are satisfied provided that* $\theta(u_L, u_R)$ *satisfies*

$$0 \le \theta(u_L, u_R) \le \Theta(u_L, u_R) \equiv \max(0, \min(1, 1 + \Gamma(u_L, u_R)), \tag{3.24}$$

where

$$\Gamma(u_L, u_R) = \begin{cases} -2\ \gamma(u_L, u_R) \dfrac{\big(-\sigma(\mathcal{U}(u_R) - \mathcal{U}(u_L)) + (\mathcal{F}(u_R) - \mathcal{F}(u_L))\big)}{|u_R - u_L|^2}, & u_L \neq u_R, \\ 0, & \textit{otherwise,} \end{cases} \tag{3.25}$$

with

$$\gamma(u_L, u_R) = \begin{cases} \dfrac{a - \min(|f'(u_L)|, |f'(u_R)|\big)}{\big(a^2 - \sigma^2(u_L, u_R)\big)}, & u_L \neq u_R, \\ 1/\big(a + |f'(u_L)|\big), & \textit{otherwise.} \end{cases} \tag{3.26}$$

Observe that if $u_L \neq u_R$, $\Gamma(u_L, u_R)$ in (3.25) is well defined under the sub-characteristic condition (2.29). Notice that

$$\gamma(u_L, u_R) = \frac{a - |f'(u_L)|}{a^2 - f'(u_L)^2} + \mathcal{O}(|u_R - u_L|) = \frac{1}{a + |f'(u_L)|} + \mathcal{O}(|u_R - u_L|), \tag{3.27}$$

hence we recover (3.26) in the limit $|u_R - u_L| \to 0$. Observe that for the pairs (u_L, u_R) satisfying the entropy inequality (2.8), we get $\Theta(u_L, u_R) = 1$ as expected so that the accuracy requirement in Lemma 2.1 can be met. Furthermore, as it is well-known that general pairs of states come with a cubic entropy rate (see for instance Godlewski-Raviart [11])

$$-\sigma(\mathcal{U}(u_R) - \mathcal{U}(u_L)) + (\mathcal{F}(u_R) - \mathcal{F}(u_L)) = \mathcal{O}(|u_R - u_L|^3), \tag{3.28}$$

we deduce

$$\Gamma(u_L, u_R) = \mathcal{O}(|u_R - u_L|). \tag{3.29}$$

Therefore, Θ is expected to stay close to unity in the smooth zones of the discrete solutions and reach ultimately 1 as $u_R \to u_L$.

Expressing the relaxation entropy pair (Φ, Ψ) in terms of the convex pair (φ^-, φ^+) according to (3.15), we first observe that the relaxation entropy jump $\mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R)$ in (3.21) equivalently reads

$$\begin{aligned} \mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R) = (a - \sigma(u_L, u_R)) \left[\varphi^+\right](\theta; u_L, u_R) \\ - (a + \sigma(u_L, u_R)) \left[\varphi^-\right](\theta; u_L, u_R), \end{aligned} \tag{3.30}$$

where we have set

$$\begin{aligned} [\varphi^-](\theta; u_L, u_R) = \varphi^-\left(r_R^{-\star}(\theta)\right) - \varphi^-\left(r_L^{-\star}(\theta)\right), \\ [\varphi^+](\theta; u_L, u_R) = \varphi^+\left(r_R^{+\star}(\theta)\right) - \varphi^+\left(r_L^{+\star}(\theta)\right) \end{aligned} \tag{3.31}$$

using the characteristic variables $r_L^{\pm\star}(\theta)$ and $r_R^{\pm\star}(\theta)$ defined in (2.36). Next, the identities (2.38) imply that for all values of $\theta \in [0,1]$:

$$\varphi^-\big(r_R^{-\star}(\theta)\big) = \varphi^-\big(r_R^-\big), \quad \varphi^-\big(r_L^{+\star}(\theta)\big) = \varphi^+\big(r_L^+\big). \tag{3.32}$$

Hence $\mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R)$ actually becomes

$$\mathcal{E}\{\mathcal{U}\}(\theta; u_L, u_R) = (a - \sigma(u_L, u_R)) \left(\varphi^+ \left(r_R^{+\star}(\theta)\right) - \varphi^+ \left(r_L^+\right)\right) - (a + \sigma(u_L, u_R)) \left(\varphi^- \left(r_R^-\right) - \varphi^- \left(r_L^{-\star}(\theta)\right)\right). \tag{3.33}$$

Further notice from definition (3.17) of the derivatives $\{\varphi^\pm\}'(r)$ that the choice of the quadratic entropy $\mathcal{U}(u) = u^2/2$ with $\mathcal{U}''(u) = 1$ yields

$$\{\varphi^\pm\}''(r) = \frac{1}{2a\big(a \pm f'(h_\pm^{-1}(r))\big)}. \tag{3.34}$$

To shed light on the forthcoming developments, we keep $\{\varphi^\pm\}''(r)$ unspecified until the end of this section. The proof of Theorem 3.5 relies on the following technical result essentially motivated by (3.33) and the convex combination (2.37) for $r_L^{-\star}(\theta)$ and $r_R^{+\star}(\theta)$, stated in Corollary 2.3.

Lemma 3.6. *For any given smooth function φ^+ and any given real number θ, the following identity holds for $r_R^{+\star}(\theta)$ defined in* (2.36)*:*

$$\begin{aligned}\varphi^+(r_R^{+\star}(\theta)) = &\Big\{\theta\varphi^+(r_R^+) + (1-\theta)\varphi^+(r_L^+)\Big\} \\ &- \theta(1-\theta) \int_0^1 \Big\{(1-\theta)\{\varphi^+\}''(r_R^+(s,\theta)) \\ &+ \theta\{\varphi^+\}''(r_L^+(s,\theta))\Big\}(1-s)ds \left(r_R^+ - r_L^+\right)^2,\end{aligned} \tag{3.35}$$

where we have set

$$r_R^+(s,\theta) = sr_R^+ + (1-s)r_R^{+\star}(\theta), \quad r_L^+(s,\theta) = sr_L^+ + (1-s)r_R^{+\star}(\theta), \quad s \in [0,1]. \tag{3.36}$$

Similarly for $r_L^{-\star}(\theta)$, we have for all θ and any given smooth function φ^-

$$\begin{aligned}\varphi^-(r_L^{-\star}(\theta)) = &\Big\{(1-\theta)\varphi^-(r_R^-) + \theta\varphi^-(r_L^-)\Big\} \\ &- \theta(1-\theta) \int_0^1 \Big\{\theta\{\varphi^-\}''(r_R^-(s,\theta)) \\ &+ (1-\theta)\{\varphi^-\}''(r_L^-(s,\theta))\Big\}(1-s)ds \left(r_R^- - r_L^-\right)^2,\end{aligned} \tag{3.37}$$

where we have defined

$$r_R^-(s,\theta) = sr_R^- + (1-s)r_L^{-\star}(\theta), \quad r_L^-(s,\theta) = sr_L^- + (1-s)r_L^{-\star}(\theta), \quad s \in [0,1]. \tag{3.38}$$

Proof. First observe the identity

$$\varphi^+(r_R^+) - \varphi^+(r_R^{+\star}(\theta)) = \{\varphi^+\}'(r_R^{+\star}(\theta))\Big(r_R^+ - r_R^{+\star}(\theta)\Big) + \int_{r_R^{+\star}(\theta)}^{r_R^+} \{\varphi^+\}''(r)\Big(r_R^+ - r\Big)dr, \tag{3.39}$$

together with

$$\varphi^+(r_L^+) - \varphi^+(r_R^{+\star}(\theta)) = \{\varphi^+\}'(r_R^{+\star}(\theta))\Big(r_L^+ - r_R^{+\star}(\theta)\Big) + \int_{r_R^{+\star}(\theta)}^{r_L^+} \{\varphi^+\}''(r)\Big(r_L^+ - r\Big)dr, \tag{3.40}$$

which yields, using the definition of $r_R^{+\star}(\theta)$ in terms of the convex decomposition (2.37) stated in Corollary 2.3,

$$\begin{aligned}\varphi^+(r_R^{+\star}(\theta)) - \Big\{\theta\varphi^+(r_R^+) + (1-\theta)\varphi^+(r_L^+)\Big\} \\ = -\theta\int_{r_R^{+\star}(\theta)}^{r_R^+} \{\varphi^+\}''(r)\Big(r_R^+ - r\Big)dr - (1-\theta)\int_{r_R^{+\star}(\theta)}^{r_L^+} \{\varphi^+\}''(r)\Big(r_L^+ - r\Big)dr.\end{aligned} \tag{3.41}$$

Introducing $r_R^+(s,\theta) = sr_R^+ + (1-s)r_R^{+\star}(\theta)$ with $s \in [0,1]$, a convenient form of the first integral in (3.41) reads

$$\begin{aligned}\int_{r_R^{+\star}(\theta)}^{r_R^+} \{\varphi^+\}''(r)\Big(r_R^+ - r\Big)dr &= \int_0^1 \{\varphi^+\}''(r_R^+(s,\theta))(1-s)ds\ \Big(r_R^+ - r_R^{+\star}(\theta)\Big)^2 \\ &= (1-\theta)^2\ \int_0^1 \{\varphi^+\}''(r_R^+(s,\theta))(1-s)ds\ \Big(r_R^+ - r_L^+\Big)^2,\end{aligned} \tag{3.42}$$

thanks again to the linear decomposition (2.37) of $r_R^{+\star}(\theta)$. Defining similarly $r_L^+(s,\theta) = sr_L^+ + (1-s)r_R^{+\star}(\theta)$ with $s \in [0,1]$, the second integral in (3.41) can be equivalently rewritten as

$$\int_{r_R^{+\star}(\theta)}^{r_L^+} \{\varphi^+\}''(r)\Big(r_L^+ - r\Big)dr = \theta^2\ \int_0^1 \{\varphi^+\}''(r_L^+(s,\theta))(1-s)ds\ \Big(r_R^+ - r_L^+\Big)^2. \tag{3.43}$$

Hence, the representation formula (3.41) becomes

$$\begin{aligned}&\varphi^+(r_R^{+\star}(\theta)) - \Big\{\theta\varphi^+(r_R^+) + (1-\theta)\varphi^+(r_L^+)\Big\} \\ &= -\theta(1-\theta)\ \int_0^1 \Big\{(1-\theta)\{\varphi^+\}''(r_R^+(s,\theta)) \\ &\qquad + \theta\{\varphi^+\}''(r_L^+(s,\theta))\Big\}(1-s)ds\ \Big(r_R^+ - r_L^+\Big)^2.\end{aligned} \tag{3.44}$$

This is nothing but the required identity (3.35). The companion formula (3.36) follows using similar steps that are left to the reader. ☐

We are now in a position to prove Theorem 3.5.

Proof of Theorem 3.5. The representation formulas (3.35) and (3.37) that are at the core of the proof, exhibit a rather intricate non-linear dependence in θ through the mappings $r_L^{\pm\star}(\theta,s)$ and $r_R^{\pm\star}(\theta,s)$ in (3.36)–(3.38). For the sake of simplicity, we aim at lowering such a dependence to a quadratic one when introducing suitable lower-bounds of the integral remainder in the Taylor-like expansions (3.35)–(3.37). For that purpose, it suffices to propose a common positive lower-bound, say $m^+(u_L,u_R)$, for $\{\varphi^+\}''(r_L^{+\star}(s,\theta))$ and $\{\varphi^+\}''(r_R^{+\star}(s,\theta))$ for all the θ and s under consideration to get

$$\begin{aligned}&\theta(1-\theta)\int_0^1 \Big\{(1-\theta)\{\varphi^+\}''(r_R^+(s,\theta)) + \theta\{\varphi^+\}''(r_L^+(s,\theta))\Big\}(1-s)ds \\ &\geq \frac{1}{2}\theta(1-\theta)m^+(u_L,u_R).\end{aligned} \tag{3.45}$$

A similar estimate holds for (3.37) adopting the same procedure with some positive lower bound $m^-(u_L,u_R)$. Here again for simplicity we adopt a common lower-bound $m(u_L, u_R) = m^-(u_L, u_R) = m^+(u_L, u_R)$. As already mentioned before, Corollary 2.3 ensures that monotonicity preserving mappings $\theta(u_L, u_R)$ make $r_L^+(s,\theta)$, $r_R^+(s,\theta)$ (respectively, $r_L^-(s,\theta)$, $r_R^-(s,\theta)$) cover the range $\lfloor r_L^+, r_R^+\rceil$ (respectively, $\lfloor r_L^-, r_R^-\rceil$) as s and θ jointly vary in $[0,1]$. Consequently, both $h_+^{-1}(r)$ and $h_-^{-1}(r)$ keep their values in $\lfloor u_L, u_R\rceil$ for all the r under consideration. The lower-bound m we seek for, must therefore satisfy

$$\min_{u\in\lfloor u_L,u_R\rceil}\left(\frac{1}{2a(a+f'(u))}, \frac{1}{2a(a-f'(u))}\right) \geq m(u_L, u_R). \tag{3.46}$$

Since the flux function f is assumed to be genuinely non-linear, the minimum in the left-hand side is achieved for $u = u_L$ or $u = u_R$ and we can thus choose

$$m(u_L, u_R) = \frac{1}{2a\big(a - \min(|f'(u_L)|, |f'(u_R)|)\big)}. \tag{3.47}$$

Plugging the proposed estimate in the representation formulas (3.35) and (3.37) immediately gives:
(3.48)

$$\varphi^+(r_R^{+\star}(\theta)) \leq \Big\{\theta\varphi^+(r_R^+) + (1-\theta)\varphi^+(r_L^+)\Big\} - \frac{\theta(1-\theta)}{2} m(u_L, u_R)\ |r_R^+ - r_L^+|^2,$$

$$\varphi^-(r_L^{-\star}(\theta)) \leq \Big\{(1-\theta)\varphi^-(r_R^-) + \theta\varphi^-(r_L^-)\Big\} - \frac{\theta(1-\theta)}{2} m(u_L, u_R)\ |r_R^- - r_L^-|^2,$$

where $|r_R^- - r_L^-| = (a-\sigma)|u_R - u_L|$ and $|r_R^+ - r_L^+| = (a+\sigma)|u_R - u_L|$. One can therefore bound the relaxation entropy jump $\mathcal{E}\{\mathcal{U}\}(\theta, u_L, u_R)$ defined in (3.33) according to

$$\begin{aligned}
\mathcal{E}\{\mathcal{U}\}(\theta, u_L, u_R) &= (a-\sigma)\Big(\varphi^+\left(r_R^{+\star}(\theta)\right) - \varphi^+\left(r_L^+\right)\Big) \\
&\quad + (a+\sigma)\Big(\varphi^-\left(r_L^{-\star}(\theta)\right) - \varphi^-\left(r_R^-\right)\Big) \\
&\leq \theta\Big\{(a-\sigma)\big(\varphi^+(r_R^+) - \varphi^+(r_L^+)\big) + (a+\sigma)\big(\varphi^-(r_R^-) - \varphi^-(r_L^-)\big)\Big\} \\
&\quad - \frac{\theta(1-\theta)}{2} m(u_L, u_R)\ \Big\{(a-\sigma)|r_R^+ - r_L^+|^2 + (a+\sigma)|r_R^- - r_L^-|^2\Big\} \\
&= \theta\Big\{-\sigma(\mathcal{U}(u_R) - \mathcal{U}(u_L)) + (\mathcal{F}(u_R) - \mathcal{F}(u_L)\Big\} \\
&\quad - \theta(1-\theta)a(a^2-\sigma^2)m(u_L, u_R)|u_R - u_L|^2.
\end{aligned} \tag{3.49}$$

Notice that, following exactly the same steps as those developed to get (3.22), we have

$$\begin{aligned}
(a-\sigma)\big(\varphi^+(r_R^+) - \varphi^+(r_L^+)\big) + (a+\sigma)\big(\varphi^-(r_R^-) - \varphi^-(r_L^-)\big) \\
= -\sigma(\mathcal{U}(u_R) - \mathcal{U}(u_L)) + (\mathcal{F}(u_R) - \mathcal{F}(u_L)) \\
= \mathcal{E}\{\mathcal{U}\}(1, u_L, u_R).
\end{aligned}$$

Then the estimate (3.49) gives

$$\mathcal{E}\{\mathcal{U}\}(\theta, u_L, u_R) \leq \theta\Big(\mathcal{E}\{\mathcal{U}\}(1, u_L, u_R) - (1-\theta)A(u_L, u_R)\Big), \tag{3.50}$$

where

$$A(u_L, u_R) = a(a^2 - \sigma^2(u_L, u_R))m(u_L, u_R)|u_R - u_L|^2. \tag{3.51}$$

Hence, assume $u_L \neq u_R$, (3.50) simply reads:

$$\mathcal{E}\{\mathcal{U}\}(\theta, u_L, u_R) \leq A(u_L, u_R)\Big\{\theta\big(\theta - (1 + \Gamma(u_L, u_R))\big)\Big\} \tag{3.52}$$

with $\Gamma(u_L, u_R)$ defined in (3.25). Since $A(u_L, u_R) > 0$, it suffices to require $\theta\big(\theta - (1+\Gamma(u_L, u_R))\big) \leq 0$ to ensure the expected entropy inequality $\mathcal{E}\{\mathcal{U}\}(\theta, u_L, u_R) \leq 0$ for the pair of states (u_L, u_R) under consideration. Enforcing as mandatory the monotonicity preserving condition $0 \leq \theta(u_L, u_R) \leq 1$ thus yields the condition (3.24). This concludes the proof.

We would like to emphasize that the upper-bound (3.50) is sharp with respect to our main motivation. Indeed, it boils down to the equality $\mathcal{E}\{\mathcal{U}\}(\theta = 1, u_L, u_R) = \mathcal{E}\{\mathcal{U}\}(1, u_L, u_R)$ and therefore it exactly preserves all the pairs (u_L, u_R) of interest, i.e., those satisfy the entropy condition $\mathcal{E}\{\mathcal{U}\}(1, u_L, u_R) \leq 0$.

3.4. **Entropy consistency for general flux functions.** To begin with, it is worth briefly recalling a few well-known facts about the Kružkov entropy criterion for selecting admissible pairs of states (u_L, u_R) that satisfy the Rankine-Hugoniot relation

$$-\sigma(u_L, u_R)(u_R - u_L) + (f(u_R) - f(u_L)) = 0. \tag{3.53}$$

The Kružkov entropy inequalities read

$$\begin{aligned} &-\sigma(u_L, u_R)\big(|u_R - k| - |u_L - k|\big) \\ &\quad + \big(\operatorname{sign}(u_R - k)(f(u_R) - f(k)) - \operatorname{sign}(u_L - k)(f(u_L) - f(k))\big) \leq 0, \end{aligned} \tag{3.54}$$

for all $k \in \mathbb{R}$. To discard empty intervals from the discussion, we tacitly assume that $u_L \neq u_R$. In (3.54), parameter k outside of the interval $\lfloor u_L, u_R \rceil$ are easily seen to satisfy the Rankine-Hugoniot jump relation (3.53), so that only the values of k in $\lfloor u_L, u_R \rceil$ are entropy diminishing

$$\operatorname{sign}(u_R - u_L)\Big\{-\sigma(u_L, u_R)\big(u_R + u_L - 2k\big) + \big(f(u_R) + f(u_L) - 2f(k)\big)\Big\} \leq 0. \tag{3.55}$$

In view of (3.53), this requirement is equivalent to the so-called Oleinik inequalities:

$$\begin{aligned} \mathcal{K}(k; u_L, u_R) &:= \operatorname{sign}(u_R - u_L)\Big\{ -\sigma(u_L, u_R)\big(u_R - k\big) + \big(f(u_R) - f(k)\big)\Big\} \\ &= \operatorname{sign}(u_R - u_L)\Big\{ -\sigma(u_L, u_R)\big(u_L - k\big) + \big(f(u_L) - f(k)\big)\Big\} \\ &\leq 0, \quad k \in \lfloor u_L, u_R \rceil. \end{aligned} \tag{3.56}$$

The main result of this section follows.

Theorem 3.7. *Let us consider the Kružkov entropy pairs* $(\mathcal{U}_k(u), \mathcal{F}_k(u))$ (3.13) *with* $k \in \lfloor u_L, u_R \rceil$ *and the associated relaxation entropy pairs* (Φ_k, Ψ_k) (3.15)–(3.16). *Assume the sub-characteristic condition* (2.29) *and consider monotonicity preserving mappings* $\theta(u_L, u_R)$ (3.2). *Then the relaxation entropy jump* $\mathcal{E}\{\mathcal{U}_k\}(\theta; u_L, u_R)$ *in* (3.21) *stay non-positive for all* $k \in \lfloor u_L, u_R \rceil$ *provided that* $\theta(u_L, u_R)$ *is chosen to satisfy*

$$0 \leq \theta(u_L, u_R) \leq \Theta(u_L, u_R) = \min_{k \in \lfloor u_L, u_R \rceil} \Big(1 + \Gamma_{\mathcal{K}}(k; u_L, u_R)\Big), \tag{3.57}$$

where

$$\Gamma_{\mathcal{K}}(k;u_L,u_R) = -2\gamma(u_L,u_R)\begin{cases} \dfrac{\mathcal{K}(k;u_L,u_R)}{|u_R-u_L|}, & \text{if } u_L \neq u_R, \\ 0, & \text{otherwise}, \end{cases} \tag{3.58}$$

with

$$\gamma(u_L,u_R) = \frac{a}{a^2-\sigma^2(u_L,u_R)}. \tag{3.59}$$

For any given pair of states (u_L,u_R), $\Theta(u_L,u_R)$ *takes value in* $(0,1)$ *and there exists at least one minimizer* $k(u_L,u_R)$ *of* $\Gamma_{\mathcal{K}}(k;u_L,u_R)$ *in* $\lfloor u_L,u_R\rceil$ *with the property that*

$$\begin{aligned} &\Theta(u_L,u_R) = 1 \text{ if } \mathcal{K}(k;u_L,u_R) \leq 0 \text{ for all } k \in \lfloor u_L,u_R\rceil, \quad \text{and} \\ &0 < \Theta(u_L,u_R) < 1 \text{ otherwise.} \end{aligned} \tag{3.60}$$

The function $\Gamma_{\mathcal{K}}(k;u_L,u_R)$ is directly built from the function $\mathcal{K}(k;u_L,u_R)$ from the Oleinik inequalities (3.56). It is easy to check that $\Gamma_{\mathcal{K}}(k;u_L,u_R) > -1$ using the sub-characteristic condition (2.29). In this respect, $\Theta(u_L,u_R)$ (3.57)–(3.58) is nothing but a natural extension of the corresponding formula (3.24)–(3.25) derived for the genuinely non-linear flux functions. Notice that in the limit $|u_R-u_L| \to 0$, we get $\Theta(u_L,u_R) = 1$. Therefore, in the numerical application, this means that the method is asymptotically close (in terms of the mesh step Δx) to a Roe solver in the smooth parts of the discrete solution.

In order to prove Theorem 3.7, we define the following two functions of the parameter k,

$$\mathcal{R}_-(k) = \frac{r_R^- - h_-(k)}{r_R^- - r_L^-}, \quad \mathcal{R}_+(k) = \frac{h_+(k) - r_L^+}{r_R^+ - r_L^+}, \quad k \in \lfloor u_L,u_R\rceil, \tag{3.61}$$

based on the characteristic variables $r^\pm = v \pm au$ and the invertible mappings $h_\pm$ defined in (3.9). Direct calculations imply

$$\begin{aligned} &\frac{a^2-\sigma^2(u_L,u_R)}{2a}(u_R-u_L)\Big(\mathcal{R}_-(k)+\mathcal{R}_+(k)\Big) \\ &\quad = \sigma(u_L,u_R)\big(u^\star(u_L,u_R)-k\big) - \big(v^\star(u_L,u_R)-f(k)\big) \\ &\quad = \frac{a^2-\sigma^2(u_L,u_R)}{2a}(u_R-u_L) + \Big(\sigma(u_L,u_R)(u_L-k) - \big(f(u_L)-f(k)\big)\Big), \end{aligned}$$

where the states $u^\star(u_L,u_R)$ and $v^\star(u_L,u_R)$ are defined in (2.30). As this formula is related to the definition of $\Gamma_{\mathcal{K}}(u_L,u_R,k)$ in (3.58), we claim that the following statement is equivalent to Theorem 3.7.

Theorem 3.8. *Under the assumptions of Theorem* 3.7, *the relaxation entropy jump* $\mathcal{E}\{\mathcal{U}_k\}(\theta;u_L,u_R)$ *in* (3.21) *stays non-positive for all* $k \in \lfloor u_L,u_R\rceil$ *provided that* $\theta(u_L,u_R)$ *is chosen to satisfy*

$$0 \leq \theta(u_L,u_R) \leq \Theta(u_L,u_R) = \min_{k\in\lfloor u_L,u_R\rceil} \Big\{\mathcal{R}_-(k)+\mathcal{R}_+(k)\Big\}, \tag{3.62}$$

where $\Theta(u_L,u_R) \in (0,1)$. *For any given pair of states* (u_L,u_R), *there exists at least one minimizer* $k(u_L,u_R)$ *of* $\mathcal{R}_-(k)+\mathcal{R}_+(k)$ *in* $\lfloor u_L,u_R\rceil$ *with the property that*

$$\begin{aligned} &\Theta(u_L,u_R) = 1 \text{ if } \mathcal{K}(k;u_L,u_R) \leq 0 \text{ for all } k \in \lfloor u_L,u_R\rceil, \quad \text{and} \\ &0 < \Theta(u_L,u_R) < 1 \text{ otherwise.} \end{aligned} \tag{3.63}$$

The proof of this theorem is postponed to the end of the section. In order to investigate the properties of the optimal choice $\Theta(u_L, u_R)$ in (3.62), we notice the following properties for the functions $\mathcal{R}_\pm(k)$. First, $\mathcal{R}_\pm(k)$ keep values in $[0,1]$ as k varies in $\lfloor u_L, u_R \rceil$ because $h_-(k)$ (respectively, $h_+(k)$) covers $\lfloor r_R^-, r_L^- \rceil$ (respectively, $\lfloor r_L^+, r_R^+ \rceil$). In addition, it is easily seen from the definitions of $h_\pm(k)$ that

$$\mathcal{R}_-(u_L) + \mathcal{R}_+(u_L) = \mathcal{R}_-(u_R) + \mathcal{R}_+(u_R) = 1. \tag{3.64}$$

Hence $\Theta(u_L, u_R) \in [0,1]$ and is thus automatically monotonicity preserving. Next, because of identity (3.64), the mapping $k \in \lfloor u_L, u_R \rceil \to \mathcal{R}_-(k) + \mathcal{R}_+(k)$ has clearly at least one extremum. As a consequence of a forthcoming representation formula for the entropy jump $\mathcal{E}\{\mathcal{U}\}_k(\theta; u_L, u_R)$, we prove that all the existing extrema stay necessarily larger than 1 in the case the pair (u_L, u_R) under consideration obeys the Kruzkov's selection principle $\mathcal{K}(k; u_L, u_R) \leq 0$ for all $k \in \lfloor u_L, u_R \rceil$. Hence we get from (3.64) the expected value $\Theta(u_L, u_R) = 1$. For other pairs, it will be seen that there exists necessarily one local minimizer $k_m(u_L, u_R)$ with the property that $(\mathcal{R}_- + \mathcal{R}_+)(k_m(u_L, u_R)) < 1$. Under the entropy inequality, one has $0 < \Theta(u_L, u_R) < 1$.

The proof of Theorem 3.8 relies on the following technical result.

Lemma 3.9. *Given a smooth enough entropy pair* $(\mathcal{U}, \mathcal{F})$ (2.2) *and the corresponding relaxation entropy pair* (Φ, Ψ) (3.15)–(3.16). *Consider monotonicity preserving mappings* $\theta(u_L, u_R)$ (3.2). *Let us define from the pair of state* (u_L, u_R) *the following affine functions of the Riemann invariants:*

$$r_-(z) = r_R^- + z(r_L^- - r_R^-), \quad r_+(z) = r_L^+ + z(r_R^+ - r_L^+), \quad z \in [0,1]. \tag{3.65}$$

Then the relaxation entropy jump $\mathcal{E}\{U\}(\theta; u_L, u_R)$ *in* (3.21) *equivalently reads*

(3.66)
$$\mathcal{E}\{U\}(\theta; u_L, u_R) = \frac{a^2 - \sigma^2(u_L, u_R)}{2a}(u_R - u_L)\int_0^\theta \Big\{\mathcal{U}'\big((h_+^{-1}(r_+(z))\big) - \mathcal{U}'\big((h_-^{-1}(r_-(z))\big)\Big\}dz.$$

Proof. Rewrite the entropy jump $\mathcal{E}\{U\}(\theta; u_L, u_R)$ for the pair (Φ, Ψ) in terms of the underlying convex pair (φ^-, φ^+) in (3.15):

(3.67)
$$\mathcal{E}\{U\}(\theta) = (a-\sigma)\Big(\varphi^+(r_R^{+\star}(\theta)) - \varphi^+(r_L^{+\star}(\theta))\Big) - (a+\sigma)\Big(\varphi^-(r_R^{-\star}(\theta)) - \varphi^-(r_L^{-\star}(\theta))\Big),$$

where by construction from Corollary 2.3, one has for all $\theta \in [0,1]$ the following convex decompositions:

$$\begin{aligned} r_L^{-\star}(\theta) &= r_R^- + \theta(r_L^- - r_R^-), & r_R^{-\star}(\theta) &= r_R^-, \\ r_L^{+\star}(\theta) &= r_L^+, & r_R^{+\star}(\theta) &= r_L^+ + \theta(r_R^+ - r_L^+). \end{aligned} \tag{3.68}$$

We can thus rewrite (3.67) as

$$\begin{aligned}\mathcal{E}\{U\}(\theta) = (a-\sigma)\int_0^1 \varphi^{+'}\big(r_L^+ + s\theta(r_R^+ - r_L^+)\big)ds\,\Big\{\theta(r_R^+ - r_L^+)\Big\} \\ + (a+\sigma)\int_0^1 \varphi^{-'}\big(r_R^- + s\theta(r_L^- - r_R^-)\big)ds\,\Big\{\theta(r_L^- - r_R^-)\Big\},\end{aligned}$$

where $r_R^+ - r_L^+ = (a+\sigma)(u_R - u_L)$ and $r_L^- - r_R^- = (a-\sigma)(u_R - u_L)$. By construction $\varphi^{\pm'}(r) = \pm U'(h_\pm^{-1}(r))/(2a)$, the above identity rewrites

$$\begin{aligned}\mathcal{E}\{U\}(\theta) &= \frac{(a^2-\sigma^2)}{2a}\big(u_R - u_L\big)\\ &\quad\times\left\{\theta\int_0^1 \mathcal{U}'\big(h_+^{-1}(r_+(s\theta))\big)\big)ds - \theta\int_0^1 \mathcal{U}'\big(h_-^{-1}(r_-(s\theta))\big)ds\right\},\end{aligned} \tag{3.69}$$

where $r_\pm$ denote the affine functions introduced in (3.65) but evaluated in $z = s\theta$. A change of variable gives the conclusion. □

In particular, we have the following result regarding the family of Kružkov entropies.

Lemma 3.10. *Consider monotonicity preserving mappings* $\theta(u_L, u_R)$ (3.2). *Then the Kružkov entropy jump* (3.54) *for any given* $k \in \lfloor u_L, u_R \rceil$ *writes*

$$\begin{aligned}&\mathcal{E}\{\mathcal{U}_k\}(\theta; u_L, u_R)\\ &\quad= -\frac{a^2-\sigma^2(u_L,u_R)}{2a}|u_R - u_L|\Big\{\mathcal{R}_-(k) + \mathcal{R}_+(k) - |\theta - \mathcal{R}_-(k)| - |\theta - \mathcal{R}_+(k)|\Big\}.\end{aligned} \tag{3.70}$$

Proof. Lemma 3.9 ensures that the relaxation entropy jump with the Kružkov entropy pair (3.13) writes

$$\begin{aligned}\mathcal{E}\{\mathcal{U}_k\}(\theta) &= \frac{a^2-\sigma^2(u_L,u_R)}{2a}|u_R - u_L|\\ &\quad\times\int_0^\theta \operatorname{sign}(u_R - u_L)\Big\{\mathcal{U}_k'\big((h_+^{-1}(r_+(z))\big) - \mathcal{U}_k'\big((h_-^{-1}(r_-(z))\big)\Big\}dz,\end{aligned} \tag{3.71}$$

where

$$\begin{aligned}\mathcal{U}_k'\big(h_-^{-1}(r_-(z))\big) &= \begin{cases} -1, & h_-^{-1}\big(r_-(z)\big) < k,\\ +1, & h_-^{-1}\big(r_-(z)\big) > k,\end{cases}\\ \mathcal{U}_k'\big(h_+^{-1}(r_+(z))\big) &= \begin{cases} -1, & h_+^{-1}\big(r_+(z)\big) < k,\\ +1, & h_+^{-1}\big(r_+(z)\big) > k.\end{cases}\end{aligned} \tag{3.72}$$

Recall that under the sub-characteristic condition (2.29), $h_-^{-1}(r)$ strictly decreases while $h_+^{-1}(r)$ strictly increases so that (3.72) reads equivalently:

$$\begin{aligned}\mathcal{U}_k'\big(h_-^{-1}(r_-(z))\big) &= \begin{cases} +1, & r_-(z) < h_-(k),\\ -1, & r_-(z) > h_-(k),\end{cases}\\ \mathcal{U}_k'\big(h_+^{-1}(r_+(z))\big) &= \begin{cases} -1, & r_+(z) < h_+(k),\\ +1, & r_+(z) > h_+(k).\end{cases}\end{aligned} \tag{3.73}$$

Easy calculations based on the sign of $(u_R - u_L)$ with $r_R^+ - r_L^+ = (a-\sigma)(u_R - u_L)$ and $r_L^- - r_R^- = (a+\sigma)(u_R - u_L)$ then allow one to recast (3.73) as

$$\begin{aligned}\operatorname{sign}(u_R - u_L)\,\mathcal{U}_k'\big(h_-^{-1}(r_-(z))\big) &= \begin{cases} +1, & z < \mathcal{R}_-(k),\\ -1, & z > \mathcal{R}_-(k),\end{cases}\\ \operatorname{sign}(u_R - u_L)\,\mathcal{U}_k'\big(h_+^{-1}(r_+(z))\big) &= \begin{cases} -1, & z < \mathcal{R}_+(k),\\ +1, & z > \mathcal{R}_+(k),\end{cases}\end{aligned} \tag{3.74}$$

where $\mathcal{R}_\pm(k)$ is defined in (3.61). Then

$$\begin{aligned}&\int_0^\theta \operatorname{sign}(u_R-u_L)\mathcal{U}_k'\big((h_-^{-1}(r_-(z))\big)dz\\ &\qquad= (+1)\min\big(\theta,\mathcal{R}_-(k)\big)+(-1)\big(\theta-\mathcal{R}_-(k)\big)_+,\\ &\qquad= \mathcal{R}_-(k)-\big(\mathcal{R}_-(k)-\theta\big)_+-\big(\theta-\mathcal{R}_-(k)\big)_+\\ &\qquad= \mathcal{R}_-(k)-|\theta-\mathcal{R}_-(k)|,\end{aligned} \tag{3.75}$$

where we have used the identity $\min(a,b)=b-\big(b-a\big)_+$ with $\big(b-a\big)_+=\max(0,b-a)$ for any given pair of real numbers (a,b). Similarly, one can infer

$$\int_0^\theta \operatorname{sign}(u_R-u_L)\mathcal{U}_k'\big((h_+^{-1}(r_+(z))\big)dz = -\Big(\mathcal{R}_+(k)-|\theta-\mathcal{R}_+(k)|\Big), \tag{3.76}$$

so that the required identity (3.70) follows from (3.71). □

As a consequence, we have the following important result.

Corollary 3.11. *Given any pair of states (u_L,u_R) obeying the Kružkov entropy condition $\mathcal{K}(k;u_L,u_R)\le 0$ for all $k\in\lfloor u_L,u_R\rceil$ in* (3.56)*. Then*

$$\min_{k\in\lfloor u_L,u_R\rceil}(\mathcal{R}_-(k)+\mathcal{R}_+(k))=1. \tag{3.77}$$

If there exists $k_\star$ in $\lfloor u_L,u_R\rceil$ with the property $\mathcal{K}(k_\star;u_L,u_R)>0$, namely the pair (u_L,u_R) is entropy violating, then

$$\min_{k\in\lfloor u_L,u_R\rceil}(\mathcal{R}_-(k)+\mathcal{R}_+(k))<1. \tag{3.78}$$

Proof. Assume an entropy satisfying pair (u_L,u_R). Then from Lemma 2.1, we have on the one hand from (3.22),

$$\mathcal{E}\{\mathcal{U}_k\}(\theta=1)=\mathcal{K}(k;u_L,u_R)\le 0 \quad \text{for all } k \text{ under consideration,} \tag{3.79}$$

while on the other hand, the representation formula (3.70) asserts that

$$\begin{aligned}&\mathcal{E}\{\mathcal{U}_k\}(\theta=1;u_L,u_R)\\ &\quad=-\frac{a^2-\sigma^2(u_L,u_R)}{2a}|u_R-u_L|\Big\{\mathcal{R}_-(k)+\mathcal{R}_+(k)-|1-\mathcal{R}_-(k)|-|1-\mathcal{R}_+(k)|\Big\}.\end{aligned} \tag{3.80}$$

Therefore, we have

$$\mathcal{R}_-(k)+\mathcal{R}_+(k)\ge|1-\mathcal{R}_-(k)|+|1-\mathcal{R}_+(k)|. \tag{3.81}$$

Since both functions $\mathcal{R}_\pm(k)$ keep their values in $[0,1]$, we get

$$2(\mathcal{R}_-(k)+\mathcal{R}_+(k))\ge 2 \quad \text{for all } k \text{ under consideration.} \tag{3.82}$$

This implies the estimate (3.77), and the upper-bound of the equality is achieved at $k=u_L$ and $k=u_R$ in view of (3.64). Next, assume there exists some $k_\star$ in $\lfloor u_L,u_R\rceil$ with the property $\mathcal{K}(k_\star;u_L,u_R)>0$. We check that any monotonicity preserving mapping $\theta(u_L,u_R)$ cannot achieve the value 1 for the pair (u_L,u_R) under consideration. Assuming there exists one such mapping then the above steps would apply to infer

$$\mathcal{R}_-(k_\star)+\mathcal{R}_+(k_\star)<|1-\mathcal{R}_-(k_\star)|+|1-\mathcal{R}_+(k_\star)|, \quad \text{i.e., } \mathcal{R}_-(k_\star)+\mathcal{R}_+(k_\star)<1, \tag{3.83}$$

and this would result in a contradiction $1 = \theta(u_L, u_R) \le \Theta(u_L, u_R) < 1$ according to the definition (3.62) of $\Theta(u_L, u_R)$. As a consequence, no monotonicity preserving mapping can reach the value 1 for the pair under consideration and we necessarily have $\Theta(u_L, u_R) < 1$. □

We conclude this section by proving Theorem 3.8.

Proof of Theorem 3.8. We first assume that the mappings $\theta(u_L, u_R)$ under consideration are monotonicity preserving and then we will prove that the resulting conditions actually imply this property. Define

$$\mathcal{H}(k, \theta) \equiv \mathcal{R}_-(k) + \mathcal{R}_+(k) - |\theta - \mathcal{R}_-(k)| - |\theta - \mathcal{R}_+(k)|. \tag{3.84}$$

In view of Lemma 3.10, limiting the values of θ such that $\mathcal{E}\{\mathcal{U}_k\}(\theta; u_L, u_R) \le 0$ for all $k \in \lfloor u_L, u_R \rceil$ is equivalent to find θ with the property

$$\mathcal{H}(k, \theta) \ge 0 \quad \text{for all } k \in \lfloor u_L, u_R \rceil. \tag{3.85}$$

The identity

$$\begin{aligned} \mathcal{H}(k, \theta) = {} & \min(\mathcal{R}_-(k), \mathcal{R}_+(k)) + \max(\mathcal{R}_-(k), \mathcal{R}_+(k)) \\ & - |\theta - \min(\mathcal{R}_-(k), \mathcal{R}_+(k))| - |\theta - \max(\mathcal{R}_-(k), \mathcal{R}_+(k))| \end{aligned} \tag{3.86}$$

then yields

$$\mathcal{H}(k, \theta) = \begin{cases} 2\theta, & 0 \le \theta \le \min(\mathcal{R}_-(k), \mathcal{R}_+(k)), \\ 2\min(\mathcal{R}_-(k), \mathcal{R}_+(k)), & \min(\mathcal{R}_-(k), \mathcal{R}_+(k)) \le \theta \\ & \quad \le \max(\mathcal{R}_-(k), \mathcal{R}_+(k)), \\ 2\big(\mathcal{R}_-(k) + \mathcal{R}_+(k) - \theta\big), & \max(\mathcal{R}_-(k), \mathcal{R}_+(k)) \le \theta. \end{cases} \tag{3.87}$$

Since by assumption $\theta \ge 0$ and $\mathcal{R}_-(k)$ and $\mathcal{R}_+(k)$ are non-negative for all $k \in \lfloor u_L, u_R \rceil$, the condition (3.85) reduces to

$$\theta \le \mathcal{R}_-(k) + \mathcal{R}_+(k) \quad \text{for all } k \in \lfloor u_L, u_R \rceil \text{ such that } \max(\mathcal{R}_-(k), \mathcal{R}_+(k)) \le \theta. \tag{3.88}$$

This condition can be easily extended to the following version:

$$\theta \le \mathcal{R}_-(k) + \mathcal{R}_+(k) \quad \text{for all } k \in \lfloor u_L, u_R \rceil \tag{3.89}$$

thanks again to the non-negativity of $\mathcal{R}_-(k)$ and $\mathcal{R}_+(k)$. Notice that condition (3.86) together with the identity (3.87) for $\theta \le \min(\mathcal{R}_-(k), \mathcal{R}_+(k))$ implies that $\theta \ge 0$; while a combination of inequality (3.89) and (3.64) suggests $\theta \le (\mathcal{R}_- + \mathcal{R}_+)(u_L) = 1$. Therefore, requiring $\mathcal{E}\{\mathcal{U}_k\}(\theta; u_L, u_R) \le 0$ for all $k \in \lfloor u_L, u_R \rceil$ in turn implies the monotonicity preserving condition (3.2). □

4. The numerical approximation schemes

This section describes first order numerical methods for approximating the Kružkov solutions of a scalar conservation law, built from the Riemann solver with defect measure correction we have derived in the first part of this paper. From now on, we assume that the monitoring mapping $\theta(u_L, u_R)$ involved in the defect measures is monotonicity preserving and consistent with the entropy requirement(s) we have put forward. Convergence of the family of approximate solutions to the Kružkov solution will be proved in the next section.

We propose hereafter two variants of finite volume methods. The first numerical method stays in the spirit of Glimm's approach and is directly built from a sequence of non-interacting Riemann solutions whose values are sampled in each cell. The second method is more in the spirit of Godunov's method and relies on suitable local averaging of two neighboring Riemann solutions. Both strategies intend to restore at the discrete level the exactness property highlighted in Lemma 2.1. To this scope, a relevant choice for the monotonicity preserving and entropy satisfying mappings θ is given by $\Theta(u_L, u_R)$, namely, either given by (3.24)–(3.26) in the case of genuinely non-linear flux function or by (3.62) for a general flux.

Introduce the spatial grid points $x_{j+\frac{1}{2}}$ with uniform mesh width $\Delta x = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$. The discrete time level t^n is also spaced uniformly with time step $\Delta t = t^{n+1} - t^n$ and satisfies the strict CFL condition

$$a\frac{\Delta t}{\Delta x} < \frac{1}{2}, \tag{4.1}$$

where the sub-characteristic condition is specified as follows:

$$\sup_{|u|<\|u_0\|_{L^\infty(\mathbb{R})}} |f'(u)| < a. \tag{4.2}$$

The numerical solution $\mathbb{U}^\alpha(t^n, x)$ is sought for as a piecewise constant function whose components are denoted by

$$u^\alpha_{\Delta x}(t^n, x) = u^n_j, \;\; v^\alpha_{\Delta x}(t^n, x) = v^n_j, \quad x_{j-\frac{1}{2}} < x < x_{j+\frac{1}{2}}, \tag{4.3}$$

where α refers to the random sequence used in Glimm's sampling procedure. The initial data is discretized in a well-prepared manner:

$$u^0_{\Delta x}(x) = \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u_0(x)dx, \quad v^0_{\Delta x}(x) = f(u^0_{\Delta x}(x)), \quad x \in (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}),\; j \in \mathbb{Z}. \tag{4.4}$$

4.1. **The first algorithm.** Assuming that the piecewise constant approximate solution $\mathbb{U}(t^n, x)$ is known at time t^n, we propose to evolve it to the next time level t^{n+1} in three steps.

$\star$ *Step* 1: $t^n \to t^{n+1,(1)} \equiv (n+1)\Delta t$, *Riemann problems with defect measure correction.* Solve the Cauchy problem exactly in the slab $(t^n, t^n + \Delta t)$:

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + a^2 \partial_x u = \mathcal{M}(u^\alpha_{\Delta x}(t^n, x), v^\alpha_{\Delta x}(t^n, x)), \end{cases} \tag{4.5}$$

with initial data

$$u(0, x) = u^\alpha_{\Delta x}(t^n, x), \;\; v(0, x) = v^\alpha_{\Delta x}(t^n, x). \tag{4.6}$$

Here $\mathcal{M}$ is a bounded Borel measure which collects all successive defect measure corrections, i.e.,

$$\begin{aligned} \mathcal{M}(u^\alpha_{\Delta x}(t^n, x), v^\alpha_{\Delta x}(t^n, x)) &= \Theta(u^n_j, u^n_{j+1})\left(a^2 - \sigma^2(u^n_j, u^n_{j+1})\right) \\ &\quad (u^n_{j+1} - u^n_j)\delta_{(x - x_{j+1/2}) - \sigma(u^n_j, u^n_{j+1})(t - t^n)} \end{aligned} \tag{4.7}$$

for $x \in (x_j, x_{j+1})$ and $t \in (t^n, t^n + \Delta t)$. Under the CFL condition (4.1), the exact solution of (4.5)–(4.6) is the gluing of a sequence of non-interacting self-similar solutions,

$$\left(\mathbb{U}^\alpha_{\Delta x}\right)^{(1)}(t, x) := \mathbb{U}\left(\frac{x - x_{j+\frac{1}{2}}}{t - t^n}; u_j, u_{j+1}\right); \quad x \in [x_j, x_{j+1}],\; t^n < t < t^{n+1}, \tag{4.8}$$

as defined in Lemma 2.2. Thus the solution at the first "intermediate" time reads

$$\left(\mathbb{U}^\alpha_{\Delta x}\right)^{(1)}\left(t^{n+1}, x\right) = \mathbb{U}\left(\frac{x - x_{j+\frac{1}{2}}}{\Delta t}; u^n_j, u^n_{j+1}\right), \;\; x \in [x_j, x_{j+1}]. \tag{4.9}$$

$\star$ *Step* 2: $t^{n+1,(1)} \to t^{n+1,(2)} \equiv (n+1)\Delta t$, *pointwise relaxation.* From the solution of Cauchy problem (4.5)–(4.6), define at the second step $t^{n+1,(2)}$ pointwisely for $x \in (x_{j-1/2}, x_{j+1/2})$:

$$u^{\alpha}_{\Delta x}{}^{(2)}\left(t^{n+1}, x\right) = u^{\alpha}_{\Delta x}{}^{(1)}\left(t^{n+1}, x\right), \tag{4.10}$$

$$v^{\alpha}_{\Delta x}{}^{(2)}\left(t^{n+1}, x\right) = f\left(u^{\alpha}_{\Delta x}{}^{(2)}\left(t^{n+1}, x\right)\right). \tag{4.11}$$

$\star$ *Step* 3: $t^{n+1,(2)} \to t^{n+1,(3)} \equiv t^{n+1}$, *sampling.* Draw a random number α_n from an equidistributed sequence in $(0,1)$, we define in each cell a constant value $\mathbb{U}^{n+1}_j$ following Glimm's sampling strategy

$$u^{\alpha}_{\Delta x}(t^{n+1}, x) = u^{\alpha}_{\Delta x}{}^{(2)}\left(t^{n+1}, x_{j-\frac{1}{2}} + \alpha_n \Delta x\right), \quad x \in [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}], \tag{4.12}$$

$$v^{\alpha}_{\Delta x}(t^{n+1}, x) = v^{\alpha}_{\Delta x}{}^{(2)}\left(t^{n+1}, x_{j-\frac{1}{2}} + \alpha_n \Delta x\right) = f(u^{\alpha}_{\Delta x}(t^{n+1}, x)). \tag{4.13}$$

This concludes the description of the method.

We summarize the first algorithm as follows.

Algorithm 4.1. *Denote*

$$\sigma^n_{j+\frac{1}{2}} = \sigma(u^n_j, u^n_{j+1}). \tag{4.14}$$

Given a random number $\alpha_n \in (0,1)$, define in each cell $(x_{j-1/2}, x_{j+1/2})$:

- *Update u^{n+1}_j from $\{u^n_j\}_{j\in\mathbb{Z}}$*

$$u^{n+1}_j = \begin{cases} u^{\star}_L(\theta; u^n_{j-1}, u^n_j), & \alpha_n < \sigma^n_{j-\frac{1}{2}} \frac{\Delta t}{\Delta x}, \\ u^{\star}_R(\theta; u^n_{j-1}, u^n_j), & \sigma^n_{j-\frac{1}{2}} \frac{\Delta t}{\Delta x} \le \alpha_n < a^n_{j-1/2} \frac{\Delta t}{\Delta x}, \\ u^n_j, & a^n_{j-\frac{1}{2}} \frac{\Delta t}{\Delta x} \le \alpha_n < 1 - a^n_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x}, \\ u^{\star}_L(\theta; u^n_j, u^n_{j+1}), & 1 - a^n_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x} \le \alpha_n < 1 + \sigma^n_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x}, \\ u^{\star}_R(\theta; u^n_j, u^n_{j+1}), & 1 + \sigma^n_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x} \le \alpha_n, \end{cases} \tag{4.15}$$

  where $u^{\star}_L$, $u^{\star}_R$ are defined in (2.31)–(2.32)*.*
- *Update $v^{n+1}_j = f(u^{n+1}_j)$.*

4.2. **The second algorithm.** Given the piecewise constant approximate solution $\mathbb{U}(t^n, x)$ at time t^n, we propose to update it to the next time level t^{n+1} in four steps. Three of these steps are virtually kept unchanged from the first numerical algorithm but are performed at (possibly) distinct intermediate stages. Here, we first summarize the basic procedure. At the first step, solve a sequence of noninteracting Riemann problems with defect measure corrections (4.5)–(4.6) to get $u^{\alpha}_{\Delta x}{}^{(1)}(t^{n+1}, x), v^{\alpha}_{\Delta x}{}^{(1)}(t^{n+1}, x)$ from $\mathbb{U}(t^n, x)$. At the second step, we perform local averaging on $u^{\alpha}_{\Delta x}{}^{(1)}(t^{n+1}, x)$ to define $u^{\alpha}_{\Delta x}{}^{(2)}(t^{n+1}, x)$. In contrast to the usual Godunov's approach, two neighboring Riemann solutions $\mathbb{U}((x - x_{j-1/2})/\Delta t, u^n_{j-1}, u^n_j)$ and $\mathbb{U}((x - x_{j+1/2})/\Delta t, u^n_j, u^n_{j+1})$ with x in $(x_{j-1/2}, x_{j+1/2})$ are not averaged within the cell under consideration. Instead, local averaging of neighboring Riemann solutions are performed over distinct intervals of the form $(x^{n+1}_{j-\frac{1}{2}}, x^{n+1}_{j+\frac{1}{2}})$ with length

$\Delta x_j^{n+1} = x_{j+\frac{1}{2}}^{n+1} - x_{j-\frac{1}{2}}^{n+1}$ and boundaries defined by

$$x_{j+\frac{1}{2}}^{n+1} = x_{j+1/2} + \sigma(u_j^n, u_{j+1}^n)\Delta t. \tag{4.16}$$

Here $x_{j+\frac{1}{2}}^{n+1}$ is the location of the intermediate discontinuity in

$$\mathbb{U}((x - x_{j+1/2})/\Delta t, u_j^n, u_{j+1}^n)$$

propagating with speed $\sigma(u_j^n, u_{j+1}^n)$ and is thus located at time $t^{n+1,(2)}$ either in $(x_{j-1/2}, x_{j+1/2})$ or in $(x_{j+1/2}, x_{j+3/2})$ depending on the sign of the velocity under consideration. The proposed local averagings are thus given by

$$u_j^{n+1,(2)} = \frac{1}{x_{j+\frac{1}{2}}^{n+1} - x_{j-\frac{1}{2}}^{n+1}} \int_{x_{j-\frac{1}{2}}^{n+1}}^{x_{j+\frac{1}{2}}^{n+1}} {u_{\Delta x}^{\alpha}}^{(1)}(t^{n+1}, x)dx, \quad j \in \mathbb{Z}. \tag{4.17}$$

This choice successfully avoids any of the intermediate waves so that it is free of numerical smearing at discontinuities. In contrast to the first algorithm, the discrete solution ${u_{\Delta x}^{\alpha}}^{(2)}(t^{n+1}, x)$ is no longer made of up to five constant states within $(x_{j-1/2}, x_{j+1/2})$ but only up to three in the situation $\sigma(u_{j-1}^n, u_j^n) > 0$ and $\sigma(u_j^n, u_{j+1}^n) < 0$. Notice that the averaging (4.17) can be wriiten in the form

$$u_j^{n+1,(2)} = \frac{\Delta x}{\Delta x_j^{n+1}} u_j^n - \frac{\Delta t}{\Delta x_j^{n+1}} \big(g_{j+1/2}^n - g_{j-1/2}^n\big), \quad j \in \mathbb{Z}, \tag{4.18}$$

where $g_{j+1/2}^n = g(u_j^n, u_{j+1}^n)$ is given by the 2-point numerical flux function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined as

$$g(u_L, u_R) = v_R^\star(\theta; u_L, u_R) - \sigma(u_L, u_R) u_R^\star(\theta; u_L, u_R), \quad (u_L, u_R) \in \mathbb{R}^2. \tag{4.19}$$

This definition leads to a conservative finite volume scheme (4.18) in view of the identity inferred from the first jump condition in (2.21)

$$v_R^\star(\theta; u_L, u_R) - \sigma(u_L, u_R) u_R^\star(\theta; u_L, u_R) = v_L^\star(\theta; u_L, u_R) - \sigma(u_L, u_R) u_L^\star(\theta; u_L, u_R). \tag{4.20}$$

Just for technical reason, the v-component is locally averaged in the second step as well, mimicking the u-component

$$v_j^{n+1,(2)} = \frac{1}{x_{j+\frac{1}{2}}^{n+1} - x_{j-\frac{1}{2}}^{n+1}} \int_{x_{j-\frac{1}{2}}^{n+1}}^{x_{j+\frac{1}{2}}^{n+1}} {v_{\Delta x}^{\alpha}}^{(1)}(t^{n+1}, x)dx, \quad j \in \mathbb{Z}. \tag{4.21}$$

At the third step, we conduct a pointwise relaxation (4.10) and get

$${v_{\Delta x}^{\alpha}}^{(3)}(t^{n+1}, x) = f({u_{\Delta x}^{\alpha}}^{(3)}(t^{n+1}, x)), \quad \text{with } {u_{\Delta x}^{\alpha}}^{(3)}(t^{n+1}, x) = {u_{\Delta x}^{\alpha}}^{(2)}(t^{n+1}, x).$$

Apparently, the third step makes the local averaging proposed for the v-component in (4.21) useless in practice. But the formal step (4.21) turns out to be convenient in the forthcoming analysis.

Within each cell $(x_{j-1/2}, x_{j+1/2})$, we derive the final update $u_{\Delta x}^{\alpha}(t^{n+1}, x)$ using a sampling procedure (4.12) performed on the piecewise constant function ${u_{\Delta x}^{\alpha}}^{(3)}(t^{n+1}, x)$. The pointwise relaxation step ensures

$$v_{\Delta x}^{\alpha}(t^{n+1}, x) = f(u_{\Delta x}^{\alpha}(t^{n+1}, x)).$$

This concludes the description of the method.

We summarize the second algorithm as follows.

Algorithm 4.2. *Given the random number $\alpha_n \in (0,1)$, in each cell $(x_{j-1/2}, x_{j+1/2})$:*

- *Update u_j^{n+1} from the $\{u_j^{n+1,(2)}\}_{j\in\mathbb{Z}}$ with $u_j^{n+1,(2)}$ given in* (4.18)

$$u_j^{n+1} = \begin{cases} u_{j-1}^{n+1,(2)}, & \alpha_n < \sigma^n_{j-\frac{1}{2}} \frac{\Delta t}{\Delta x}, \\ u_{j}^{n+1,(2)}, & \sigma^n_{j-\frac{1}{2}} \frac{\Delta t}{\Delta x} \le \alpha_n < 1 + \sigma^n_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x}, \\ u_{j+1}^{n+1,(2)}, & 1 + \sigma^n_{j+\frac{1}{2}} \frac{\Delta t}{\Delta x} \le \alpha_n. \end{cases} \tag{4.22}$$

- *Update $v_j^{n+1} = f(u_j^{n+1})$.*

5. Convergence to the Kružkov entropy weak solution

In this section, we prove for both the finite volume methods Algorithm 4.1 and 4.2 that the family of discrete solutions $\{\mathbb{U}^\alpha_{\Delta x}\}_{\Delta x>0}$ converges as Δx goes to zero to $\mathbb{U} = \big(u, f(u)\big)$ where u is the Kružkov solution of the Cauchy problem for (2.1) with initial data $u_0 \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$. The main result is as follows.

Theorem 5.1. *Given $u_0 \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$. Assume the sub-characteristic condition* (4.2) *and the CFL condition* (4.1)*. Assume that the mapping $\theta(u_L, u_R)$ is monotonicity preserving* (3.2) *and consistent with the entropy condition* (3.21)*, namely with the quadratic entropy pair in the case of a genuinely non-linear flux and with the whole Kružkov family in the case of a general non-linear flux function. Then for almost any given sampling sequence $\alpha = (\alpha_1, \alpha_2, \dots) \in (0,1)^{\mathbb{N}}$, the family of approximate solutions $\{u^\alpha_{\Delta x}\}_{\Delta x>0}$ given either by* (4.1) *or* (4.2) *converges in $L^\infty\big((0,T), L^1_{loc}(\mathbb{R})\big)$ for all $T>0$ and a.e. as $\Delta x \to 0$ with $\frac{\Delta t}{\Delta x}$ kept fixed, to the Kružkov solution of the corresponding Cauchy problem* (2.1)*.*

The proof of this statement first relies on the following result.

Proposition 5.2. *Assume the sub-characteristic condition* (4.2) *and the CFL condition* (4.1)*. Suppose that the mapping $\theta(u_L, u_R)$ is monotonicity preserving, then for any given sampling sequence $\alpha = (\alpha_1, \alpha_2, \dots) \in (0,1)^{\mathbb{N}}$, the sequence of discrete solutions $(u^\alpha_{\Delta x}(t,x), v^\alpha_{\Delta x}(t,x))_{\Delta t>0}$ obtained either by Algorithm* 4.1 *or Algorithm* 4.2 *satisfies the following uniform in Δx a priori estimates for all time $t>0$:*

$$\text{(i)} \quad \| u^\alpha_{\Delta x}(t,\cdot)\|_{L^\infty(\mathbb{R})} \le \| u_0 \|_{L^\infty(\mathbb{R})}, \quad \| v^\alpha_{\Delta x}(t,\cdot)\|_{L^\infty(\mathbb{R})} \le a\, \| u_0 \|_{L^\infty(\mathbb{R})}, \tag{5.1}$$

$$\text{(ii)} \quad TV(u^\alpha_{\Delta x}(t,.)) \le TV(u_0), \quad TV(v^\alpha_{\Delta x}(t,.)) \le a\, TV(u_0), \tag{5.2}$$

$$\text{(iii)} \quad \int_{\mathbb{R}} \left| {u^\alpha_{\Delta x}}^{(1)}(t,x) - u^\alpha_{\Delta x}(t^n,x) \right| dx \le a\, TV(u_0)(t-t^n), \quad t^n \le t \le t^{n+1}, \tag{5.3}$$

$$\text{(iv)} \int_{\mathbb{R}} \left| {v^\alpha_{\Delta x}}^{(1)}(t,x) - f(u^\alpha_{\Delta x}(t^n,x)) \right| dx \le 2a^2\, TV(u_0)(t-t^n), \; t^n \le t \le t^{n+1}. \tag{5.4}$$

Proof. We will prove the results Algorithm 4.2. The proof for (the simpler) Algorithm 4.1 follows from essentially identical steps and the details are left to the reader.

(i) The sup-norm estimate in (5.1) follows from the corresponding local maximum principle stated in (3.1), Theorem 3.1, which is valid in the first step:

(5.5)
$$\sup_{x_j\le x\le x_{j+1}} |{u_{\Delta x}^{\alpha}}^{(1)}(t,x)| = \sup_{x_j\le x\le x_{j+1}} \left| u\left(\frac{x-x_{j+\frac12}}{t-t^n}; u_j^n, u_{j+1}^n\right)\right| \le \max(|u_j^n|, |u_{j+1}^n|),$$

for all $j\in\mathbb{Z}$ and $t\in(t^n,t^{n+1})$, and as a consequence

$$\| {u_{\Delta x}^{\alpha}}^{(1)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} \le \| u_{\Delta x}^{\alpha}(t^n,\cdot)\|_{L^\infty(\mathbb{R})} \,. \tag{5.6}$$

As is well-known, the local averagings involved in the second step diminish the sup-norm

$$\| {u_{\Delta x}^{\alpha}}^{(2)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} \le \| {u_{\Delta x}^{\alpha}}^{(1)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} \,. \tag{5.7}$$

The third step devoted to pointwise relaxation does not change the u-component of the discrete solution, and the sampling procedure in the last step decreases the sup-norm, so that

$$\begin{aligned}\| u_{\Delta x}^{\alpha}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} &\le \| {u_{\Delta x}^{\alpha}}^{(3)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})}\\ &\le \| {u_{\Delta x}^{\alpha}}^{(2)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} \le \| u_{\Delta x}^{\alpha}\left(t^{n},\cdot\right)\|_{L^\infty(\mathbb{R})} \,.\end{aligned} \tag{5.8}$$

This immediately implies the expected uniform sup-norm estimate in view of the definition (4.4) of the discrete initial data. The derivation of the companion sup-norm estimate for $v_{\Delta x}^{\alpha}(t,\cdot)$ starts from the local estimate (3.4),

$$\begin{aligned}\sup_{x_j\le x\le x_{j+1}} |{v_{\Delta x}^{\alpha}}^{(1)}(t,x)| &= \sup_{x_j\le x\le x_{j+1}} \left| v\left(\frac{x-x_{j+\frac12}}{t-t^n}; u_j^n, u_{j+1}^n\right)\right|\\ &\le 2a\max(|u_j^n|, |u_{j+1}^n|)\end{aligned} \tag{5.9}$$

for all $j\in\mathbb{Z}$ and $t\in(t^n,t^{n+1})$, so that

$$\| {v_{\Delta x}^{\alpha}}^{(1)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} \le 2a\, \| u_{\Delta x}^{\alpha}(t^n,\cdot)\|_{L^\infty(\mathbb{R})} \,. \tag{5.10}$$

Then in the third step, $v_{\Delta x}^{\alpha}$ is set at equilibrium pointwisely in x, and we get from estimate (5.6):

(5.11)
$$\| {v_{\Delta x}^{\alpha}}^{(3)}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} = \| f({u_{\Delta x}^{\alpha}}^{(2)}\left(t^{n+1},\cdot\right))\|_{L^\infty(\mathbb{R})} \le 2a\, \| u_{\Delta x}^{\alpha}\left(t^{n},\cdot\right))\|_{L^\infty(\mathbb{R})} \,.$$

At last the sampling procedure does not increase the sup-norm of $v_{\Delta x}^{\alpha}$ so that

$$\| v_{\Delta x}^{\alpha}\left(t^{n+1},\cdot\right)\|_{L^\infty(\mathbb{R})} \le 2a\, \| u_{\Delta x}^{\alpha}\left(t^{n},\cdot\right))\|_{L^\infty(\mathbb{R})} \le 2a\, \| u_0\|_{L^\infty(\mathbb{R})} \,. \tag{5.12}$$

(ii) In view of the local total variation estimate stated in (3.1), the first step gives

(5.13)
$$\mathrm{TV}_{(x_j,x_{j+1})}\left({u_{\Delta x}^{\alpha}}^{(1)}(t,\cdot)\right) = \mathrm{TV}\left(u(\cdot;u_j^n,u_{j+1}^n)\right) \le |u_{j+1}^n - u_j^n|, \quad t^n\le t\le t^{n+1}.$$

Under the CFL condition (4.1), the discrete solution $u_{\Delta x}^{\alpha}(t,x)$ remains continuous at $x=x_j$ while keeping the constant value u_j^n for all $t\in(t^n,t^{n+1,(1)})$, we infer

$$\begin{aligned}\mathrm{TV}\left({u_{\Delta x}^{\alpha}}^{(1)}(t,\cdot)\right) &= \sum_{j\in\mathbb{Z}} \mathrm{TV}_{(x_j,x_{j+1})}\left({u_{\Delta x}^{\alpha}}^{(1)}(t,\cdot)\right)\\ &\le \sum_{j\in\mathbb{Z}} \left|u_{j+1}^n - u_j^n\right| = \mathrm{TV}\left(u_{\Delta x}^{\alpha}(t^n,\cdot)\right).\end{aligned} \tag{5.14}$$

In the second step, $u^\alpha_{\Delta x}$ is locally averaged and its total variation decreases:

$$\mathrm{TV}\left({u^\alpha_{\Delta x}}^{(2)}(t^{n+1},\cdot)\right) \le \mathrm{TV}\left({u^\alpha_{\Delta x}}^{(1)}(t^{n+1},\cdot)\right) \le \mathrm{TV}\left(u^\alpha_{\Delta x}(t^n,\cdot)\right). \tag{5.15}$$

In the third step, $u^\alpha_{\Delta x}$ is kept unchanged and at last, the sampling procedure clearly diminishes the total variation, thus an immediate recursion gives the required uniform total variation estimate again from the definition (4.4) of the discrete initial data

$$\mathrm{TV}(u^\alpha_{\Delta x}(t^{n+1},\cdot)) \le \mathrm{TV}\left(u^\alpha_{\Delta x}(t^n,\cdot)\right) \le \mathrm{TV}(u^0_{\Delta x}) \le \mathrm{TV}(u_0). \tag{5.16}$$

The estimate for $v^\alpha_{\Delta x}(t,\cdot)$ is derived similarly starting from the local estimate (3.4) for each self-similar solution to infer

$$\mathrm{TV}_{(x_j,x_{j+1})}\left({v^\alpha_{\Delta x}}^{(1)}(t,\cdot)\right) \le a\mathrm{TV}\left(u(\cdot;u^n_j,u^n_{j+1})\right), \quad t^n \le t \le t^{n+1}, \tag{5.17}$$

so that

$$\mathrm{TV}\left({v^\alpha_{\Delta x}}^{(1)}(t,\cdot)\right) \le a\mathrm{TV}\left(u^\alpha_{\Delta x}(t^n,\cdot)\right), \quad t^n \le t \le t^{n+1}. \tag{5.18}$$

In the second step, $v^\alpha_{\Delta x}$ is locally averaged according to (4.21) hence

$$\mathrm{TV}\left({v^\alpha_{\Delta x}}^{(2)}(t,\cdot)\right) \le a\mathrm{TV}\left(u^\alpha_{\Delta x}(t^n,\cdot)\right), \quad t^n \le t \le t^{n+1}, \tag{5.19}$$

and is then set at equilibrium in the third step

$$\mathrm{TV}\left({v^\alpha_{\Delta x}}^{(3)}(t^{n+1},\cdot)\right) = \mathrm{TV}\left(f({u^\alpha_{\Delta x}}^{(2)}(t^{n+1},\cdot))\right) \le a\mathrm{TV}\left(u^\alpha_{\Delta x}(t^n,\cdot)\right). \tag{5.20}$$

At last, the sampling procedure diminishes the total variation

$$\mathrm{TV}(v^\alpha_{\Delta x}(t^{n+1},\cdot)) \le a\mathrm{TV}\left(u^\alpha_{\Delta x}(t^n,\cdot)\right) \le a\mathrm{TV}(u_0). \tag{5.21}$$

(iii) Observe from the first step, the following identity which holds in the sense of the Radon measures

$$\partial_t {u^\alpha_{\Delta x}}^{(1)}(t,x) = -\partial_x {v^\alpha_{\Delta x}}^{(1)}(t,x), \quad t \in (t^n, t^{n+1}). \tag{5.22}$$

Under the CFL condition (4.1), the total variation of the Radon measure $\partial_t u^\alpha_{\Delta x}$ can be bounded from above by

$$|\partial_t {u^\alpha_{\Delta x}}^{(1)}(t,x)|(\mathbb{R}_x) = \mathrm{TV}({v^\alpha_{\Delta x}}^{(1)}(t,.)) \le a\mathrm{TV}(u_0) \tag{5.23}$$

so that, one can infer for $t \in (t^n, t^{n+1})$ that

$$\int_{\mathbb{R}_x} \left|{u^\alpha_{\Delta x}}^{(1)}(t,x) - u^\alpha_{\Delta x}(t^n,x)\right| dx \quad = \int_{t^n}^t |\partial_t {u^\alpha_{\Delta x}}^{(1)}(s,x)|(\mathbb{R}_x) ds \tag{5.24}$$

$$\le aTV(u_0)(t-t_n). \tag{5.25}$$

(iv) The equation involving the defect measure correction reads for $t \in (t^n, t^{n+1})$ and $x \in (x_j, x_{j+1})$ that

$$\partial_t {v^\alpha_{\Delta x}}^{(1)} = -a^2 \partial_x {u^\alpha_{\Delta x}}^{(1)} + m(u^n_j, u^n_{j+1})\delta_{x - x_{j+1/2} - \sigma(u^n_j,u^n_{j+1})(t-t^n)}, \tag{5.26}$$

and the quantities involved in the above identity are again regarded as Radon measures. The total variation of the Radon measure $\partial_t v^\alpha_{\Delta x}$ can be bounded by

$$\begin{aligned} |\partial_t {v^\alpha_{\Delta x}}^{(1)}(t,x)|(x_j,x_{j+1}) &\le a^2 |\partial_x {u^\alpha_{\Delta x}}^{(1)}|(x_j,x_{j+1}) + |m(u^n_j,u^n_{j+1})| \\ &\le a^2|u^n_{j+1} - u^n_j| + (a^2-\sigma^2)\left|u^n_{j+1}-u^n_j\right| \le 2a^2\left|u^n_{j+1}-u^n_j\right|. \end{aligned} \tag{5.27}$$

Therefore by summation, (5.27) becomes, under the CFL condition (4.1),

$$|\partial_t {v^\alpha_{\Delta x}}^{(1)}(t,x)|(\mathbb{R}_x) \le 2a^2 \mathrm{TV}\left({u^\alpha_{\Delta x}}^{(1)}(t^n,\cdot)\right) \le 2a^2\mathrm{TV}(u_0). \tag{5.28}$$

We deduce for $t^n \le t \le t^{n+1}$ that

$$\int_{\mathbb{R}_x} \left|{v^\alpha_{\Delta x}}^{(1)}(t,x) - v^\alpha_{\Delta x}(t^n,x)\right| dx = \int_{t^n}^t |\partial_t {v^\alpha_{\Delta x}}^{(1)}(s,x)|(\mathbb{R}_x)ds \le 2a^2\mathrm{TV}(u_0)(t-t^n),$$

where by construction $v^\alpha_{\Delta x}(t^n,x) = f(u^\alpha_{\Delta x}(t^n,x))$. This concludes the proof. □

This proposition immediately implies the following convergence result.

Corollary 5.3. *Given $u_0 \in L^\infty \cap BV(\mathbb{R})$, any $T>0$, then under the assumptions of Proposition* 5.2*, there exists a subsequence still denoted by $\{u^\alpha_{\Delta x}\}_{\Delta x>0}$, which converges, as $\Delta x \to 0$ with $\Delta t/\Delta x$ kept constant, to a limit u^α in $L^\infty\big((0,T), L^1_{loc}(\mathbb{R})\big)$. In addition, the limit u^α belongs to $L^\infty(\mathbb{R}_+, L^\infty \cap BV(\mathbb{R}))$.*

Proof. This proof is rather classical from the uniform estimates stated in Proposition 5.2, and one can refer for instance to [11] (Theorems 3.3 and 3.4, Chapter 3). □

The above corollary guarantees the existence of a limit. We now characterize this limit, showing that it is indeed the unique entropy weak solution of the original Cauchy problem (2.1). The proof mainly relies on the relaxation entropy inequalities inherited from the first step shared by both Algorithm 4.1 and Algorithm 4.2. Consider the following time-space domains:

$$\mathcal{D}^n_j = \Big\{(t,x)\in\mathbb{R}^+\times\mathbb{R}/\ t\in(t^n,t^{n+1}),\quad x^n_{j-1/2}(t) < x < x^n_{j+1/2}(t),\quad x^n_{j+1/2}(t) = x_{j+1/2} + \sigma^n_{j+1/2}(t-t^n)\Big\}. \tag{5.29}$$

Notice that $x^n_{j+1/2}(t^{n+1})$ coincides with $x^{n+1}_{j+1/2}$ defined in (4.16). We state the following.

Lemma 5.4. *Under the assumptions of Proposition* 5.2*, the approximate solutions given in the first step either by Algorithm* 4.1 *or Algorithm* 4.2 *satisfy the following relaxation entropy equalities in the sense of the distributions:*

$$\partial_t\Phi({u^\alpha_{\Delta x}}^{(1)}, {v^\alpha_{\Delta x}}^{(1)}) + \partial_x\Psi({u^\alpha_{\Delta x}}^{(1)}, {v^\alpha_{\Delta x}}^{(1)}) = 0, \quad (t,x)\in\mathcal{D}^n_j, n\ge 0, j\in\mathbb{Z}. \tag{5.30}$$

Assume in addition that the mapping $\theta(u_L,u_R)$ is consistent with the entropy condition (3.21)*, namely with the quadratic entropy pair in the case of a genuinely non-linear flux or with the whole Kružkov entropy family in the case of a general non-linear flux function. Then the discrete solutions given either by Algorithm* 4.1 *or Algorithm* 4.2 *satisfy the corresponding entropy jump(s) at each boundary $x^n_{j+1/2}(t)$:*

$$\begin{aligned}&-\sigma^n_{j+1/2}\Big(\Phi({u^\alpha_{\Delta x}}^{(1)}, {v^\alpha_{\Delta x}}^{(1)})(t,x^n_{j+1/2}(t)_+) - \Phi({u^\alpha_{\Delta x}}^{(1)}, {v^\alpha_{\Delta x}}^{(1)})(t,x^n_{j+1/2}(t)_-)\Big)\\ &+\Big(\Psi({u^\alpha_{\Delta x}}^{(1)}, {v^\alpha_{\Delta x}}^{(1)})(t,x^n_{j+1/2}(t)_+) - \Psi({u^\alpha_{\Delta x}}^{(1)}, {v^\alpha_{\Delta x}}^{(1)})(t,x^n_{j+1/2}(t)_-)\Big) \le 0,\\ &\qquad t\in(t^n,t^{n+1}).\end{aligned} \tag{5.31}$$

Proof. Under the strict CFL condition (4.1), two neighboring Riemann solutions do not interact. We thus observe from the definition of each of the domain $\mathcal{D}_j^n$ that the solution $(u_{\Delta x}^{\alpha}, v_{\Delta x}^{\alpha})$ is locally made of three constant states separated by the discontinuity lines $x_{j-1/2} + a(t - t^n)$ and $x_{j+1/2} - a(t - t^n)$. The property that the relaxation entropy is preserved across these two discontinuities (see indeed (3.20)) yields the expected equality (5.30). Next and for the mapping $\theta(u_L, u_R)$ under consideration, the jump inequality across each of the boundary $x_{j+1/2}^n(t)$ reads nothing but our entropy consistency requirement (3.21) stated in Definition 3.3. $\square$

As a consequence, we get the following.

Proposition 5.5. *Assume the sub-characteristic condition* (4.2) *and the CFL condition* (4.1)*. Given an entropy pair* $(\mathcal{U}, \mathcal{F})$ (2.2) *with* $\mathcal{U}$ *convex and its corresponding relaxation entropy pair* (Φ, Ψ) (3.15)–(3.16)*. Assume that the mapping* $\theta(u_L, u_R)$ *is consistent with the entropy requirement* (3.21) *for all the pairs* (u_L, u_R) *under consideration. Then for any non-negative test function* $\zeta \in C_0^1((0, \infty) \times \mathbb{R}_x)$*, the approximate solutions* $(u_{\Delta x}^{\alpha}, v_{\Delta x}^{\alpha})$ *given either by Algorithm* 4.1 *or Algorithm* 4.2 *satisfy*

$$\int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \mathcal{U}\left(u_{\Delta x}^{\alpha}(t^{n+1}, x)\right)\zeta(t^{n+1}, x)dx - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathcal{U}\left(u_{\Delta x}^{\alpha}(t^{n}, x)\right)\zeta(t^{n}, x)dx$$

$$+\mathcal{G}_{j+1/2}^n - \mathcal{G}_{j-1/2}^n - \iint_{\mathcal{D}_j^n} \Phi(u_{\Delta x}^{\alpha}, v_{\Delta x}^{\alpha})\partial_t\zeta + \Psi(u_{\Delta x}^{\alpha}, v_{\Delta x}^{\alpha})\partial_x\zeta dtdx$$

$$\leq \left(\mathcal{E}_A\right)_j^n(\Delta x, \alpha, \zeta) + \left(\mathcal{E}_S\right)_j^n(\Delta x, \alpha, \zeta), \quad n \geq 0,\ j \in \mathbb{Z}. \tag{5.32}$$

Here, $\mathcal{G}_{j+1/2}^n$ *stands for the time average of the right trace of the entropy flux along the boundary* $x_{j+1/2}^n(t)$ *and reads for both methods:*

$$\mathcal{G}_{j+1/2}^n := \int_{t^n}^{t^{n+1}} \left\{\Psi({u_{\Delta x}^{\alpha}}^{(1)}, {v_{\Delta x}^{\alpha}}^{(1)}) - \sigma_{j+1/2}^n \Phi({u_{\Delta x}^{\alpha}}^{(1)}, v_{\Delta x}^{\alpha}(1))\right\} \times (t, x_{j+\frac{1}{2}}^n(t)_+)\zeta(t, x_{j+\frac{1}{2}}^n(t))dt. \tag{5.33}$$

Concerning Algorithm 4.1*, the error term* $(\mathcal{E}_A)$ *due to local averagings is identically zero while the error term* $(\mathcal{E}_S)$ *due to the sampling procedure is given by*

$$\left(\mathcal{E}_S\right)_j^n(\Delta x, \alpha, \zeta) := \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \left(\mathcal{U}\left(u_{\Delta x}^{\alpha}(t^{n+1}, x)\right) - \mathcal{U}\left({u_{\Delta x}^{\alpha}}^{(2)}(t^{n+1}, x)\right)\right)\zeta(t^{n+1}, x)dx. \tag{5.34}$$

For Algorithm 4.2*, the error terms* $(\mathcal{E}_A)$ *and* $(\mathcal{E}_S)$*, respectively, read*

$$\left(\mathcal{E}_A\right)_j^n(\Delta x, \alpha, \zeta) := \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \left(\Phi\left(({u_{\Delta x}^{\alpha}}^{(2)}, {v_{\Delta x}^{\alpha}}^{(2)})(t^{n+1}, x)\right) - \Phi\left(({u_{\Delta x}^{\alpha}}^{(1)}, {v_{\Delta x}^{\alpha}}^{(1)})(t^{n+1}, x)\right)\right)\zeta(t^{n+1}, x)dx \tag{5.35}$$

and

$$\left(\mathcal{E}_S\right)_j^n(\Delta x, \alpha, \zeta) := \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \left(\mathcal{U}\left(u_{\Delta x}^{\alpha}(t^{n+1}, x)\right) - \mathcal{U}\left({u_{\Delta x}^{\alpha}}^{(3)}(t^{n+1}, x)\right)\right)\zeta(t^{n+1}, x)dx. \tag{5.36}$$

Remark 5.6. In (5.32), the superscript (1) has been omitted in the notation of the (volume) integral over $\mathcal{D}_j^n$ since time discontinuities in the subsequent steps $t^{n+1,(1)}$, $t^{n+1,(2)}$, $t^{n+1,(3)}$ form a negligible set in the proposed Lebesgue integral.

Proof. The proposed inequality is proved for the second algorithm in section 4.2. Its derivation for the method in section 4.1 follows the same lines. Since again $(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)$ is nothing but a piecewise constant solution of the entropy conservation law (5.30) over the domain $\mathcal{D}_j^n$, multiplying (5.30) by any given non-negative test function $\zeta \in C_0^1(\mathbb{R}_t^+ \times \mathbb{R})$ and integrating over $(t,x) \in \mathcal{D}_j^n$ yield

$$
\begin{aligned}
&\int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})\left(t^{n+1}, x\right)\zeta(t^{n+1},x)dx \\
&\quad - \int_{x_{j-\frac12}}^{x_{j+\frac12}} \Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)(t^n,x)\zeta(t^n,x)dx \\
&\quad + \int_{t^n}^{t^{n+1}} \Big\{\Psi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)}) - \sigma_{j+1/2}^n \Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})\Big\} \\
&\qquad\qquad \times (t, x_{j+1/2}^n(t)_-)\zeta(t, x_{j+1/2}^n(t))dt \\
&\quad - \int_{t^n}^{t^{n+1}} \Big\{\Psi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)}) - \sigma_{j-1/2}^n \Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})\Big\} \\
&\qquad\qquad \times \zeta(t, x_{j-1/2}^n(t)_+)\zeta(t, x_{j-1/2}^n(t))dt \\
&\quad - \iint_{\mathcal{D}_j^n} \Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_t\zeta + \Psi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_x\zeta dtdx = 0,
\end{aligned} \tag{5.37}
$$

where the left and right traces at any given interface $x_{j+1/2}^n(t)$ are well defined since both $u_{\Delta x}^\alpha(t,.)$ and $v_{\Delta x}^\alpha(t,.)$ have uniformly bounded total variation in space. Using the definition (5.33) of $\mathcal{G}_{j+1/2}^n$ evaluated on the right trace, inequality (5.37) can be recast as

$$
\begin{aligned}
&\int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})\left(t^{n+1}, x\right)\zeta(t^{n+1},x)dx \\
&\quad - \int_{x_{j-\frac12}}^{x_{j+\frac12}} \Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)(t^n,x)\zeta(t^n,x)dx + \mathcal{G}_{j+1/2}^n - \mathcal{G}_{j-1/2}^n \\
&\quad - \iint_{\mathcal{D}_j^n} \Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_t\zeta + \Psi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_x\zeta dtdx = \mathcal{S}_j^n,
\end{aligned} \tag{5.38}
$$

where

$$
\begin{aligned}
\mathcal{S}_j^n := \int_{t^n}^{t^{n+1}} &\Big\{ - \sigma_{j+1/2}^n \Big(\Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})(t, x_{j+1/2}^n(t)_+) \\
&\qquad - \Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})(t, x_{j+1/2}^n(t)_-)\Big) \\
&+ \Big(\Psi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})(t, x_{j+1/2}^n(t)_+) \\
&\qquad - \Psi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)})(t, x_{j+1/2}^n(t)_-)\Big)\Big\}\zeta(t, x_{j+\frac12}^n(t))dt \\
&\le 0.
\end{aligned} \tag{5.39}
$$

Due to the jump inequality (5.31) established in Lemma 5.4, $\mathcal{S}_j^n$ is non-positive. Then by construction $v_{\Delta x}^\alpha(t^n, x) = f(u_{\Delta x}^\alpha(t^n, x))$ for all x in $(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ so that the consistency condition (3.14) which links the entropy $\mathcal{U}$ to its relaxation extension Φ gives

$$\Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)(t^n, x) = \mathcal{U}(u_{\Delta x}^\alpha(t^n, x)), \quad x \in (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}), \ j \in \mathbb{Z}. \tag{5.40}$$

Hence, by (5.38)–(5.39), we infer

$$\begin{aligned}
&\int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \Phi({u_{\Delta x}^\alpha}^{(1)}, {v_{\Delta x}^\alpha}^{(1)}) \left(t^{n+1}, x\right) \zeta(t^{n+1}, x) dx \\
&\qquad - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathcal{U}(u_{\Delta x}^\alpha(t^n, x))\zeta(t^n, x)dx + \mathcal{G}_{j+1/2}^n - \mathcal{G}_{j-1/2}^n \\
&\qquad - \iint_{\mathcal{D}_j^n} \Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_t \zeta + \Psi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_x \zeta dt dx \le 0.
\end{aligned} \tag{5.41}$$

After the second step on the local averaging (4.17)–(4.21), we thus deduce from (5.41) the inequality

$$\begin{aligned}
&\int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \Phi({u_{\Delta x}^\alpha}^{(2)}, {v_{\Delta x}^\alpha}^{(2)})(t^{n+1}, x)\zeta(t^{n+1}, x)dx \\
&\qquad - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathcal{U}(u_{\Delta x}^\alpha(t^n, x))\zeta(t^n, x)dx + \mathcal{G}_{j+1/2}^n - \mathcal{G}_{j+1/2}^n \\
&\qquad - \iint_{\mathcal{D}_j^n} \Phi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_t \zeta + \Psi(u_{\Delta x}^\alpha, v_{\Delta x}^\alpha)\partial_x \zeta dt dx \\
&\quad \le \left(\mathcal{E}_A\right)_j^n (\Delta x, \alpha, \zeta),
\end{aligned} \tag{5.42}$$

where $\left(\mathcal{E}_A\right)_j^n (\Delta x, \alpha, \zeta)$ denotes the local averaging error term defined in (5.35). Under the sub-characteristic condition (4.2), the Gibbs principle (3.23) established in Lemma 3.4 ensures that in the third step, the following inequality holds pointwisely in x,

$$\begin{aligned}
\mathcal{U}({u_{\Delta x}^\alpha}^{(3)}) \left(t^{n+1}, x\right) &= \Phi\left({u_{\Delta x}^\alpha}^{(3)}, f({u_{\Delta x}^\alpha}^{(3)})\right) \left(t^{n+1}, x\right) \\
&\le \Phi({u_{\Delta x}^\alpha}^{(2)}, {v_{\Delta x}^\alpha}^{(2)}) \left(t^{n+1}, x\right),
\end{aligned} \tag{5.43}$$

where by construction ${u_{\Delta x}^\alpha}^{(3)}(t^{n+1}, x) = {u_{\Delta x}^\alpha}^{(2)}(t^{n+1}, x)$. The expected inequality (5.32) then holds at the end of the last step devoted to the sampling procedure, with an additional error term given by (5.36). This concludes the proof. □

We are in a position to prove the convergence of the family of discrete solutions given either by (4.1) or (4.2) to the unique Kružkov solution of (2.1).

Proof of Theorem 5.1. For any given non-negative test function $\zeta \in C_0^1\left((0,\infty)\times\mathbb{R}_x\right)$, we sum up the inequalities (5.32) for $j \in \mathbb{Z}$ to get

$$\begin{aligned}&\int_{\mathbb{R}} \mathcal{U}(u^\alpha_{\Delta x})\left(t^{n+1},x\right)\zeta(t^{n+1},x)dx - \int_{\mathbb{R}} \mathcal{U}(u^\alpha_{\Delta x})(t^n,x)\zeta(t^n,x)dx \\ &\quad - \int_{t^n}^{t^{n+1}}\int_{\mathbb{R}} \Phi(u^\alpha_{\Delta x}, v^\alpha_{\Delta x})\partial_t\zeta + \Psi(u^\alpha_{\Delta x}, v^\alpha_{\Delta x})\partial_x\zeta dtdx \\ &\quad \le \sum_{j\in\mathbb{Z}}\Big(\left(\mathcal{E}_A\right)^n_j(\Delta x,\alpha,\zeta) + \left(\mathcal{E}_S\right)^n_j(\Delta x,\alpha,\zeta)\Big), \quad n \ge 0.\end{aligned} \tag{5.44}$$

Summing with respect to $n \in \mathbb{N}$ then yields

$$\begin{aligned}&-\sum_{n\ge 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}} \Phi(u^\alpha_{\Delta x}, v^\alpha_{\Delta x})\partial_t\zeta + \Psi(u^\alpha_{\Delta x}, v^\alpha_{\Delta x})\partial_x\zeta dtdx \\ &\quad - \int_{\mathbb{R}} \mathcal{U}(u^\alpha_{\Delta x})(0,x)\zeta(0,x)dx \le \mathcal{E}_A + \mathcal{E}_S,\end{aligned} \tag{5.45}$$

where we have set

$$\mathcal{E}_A = \sum_{n\ge 0}\int_{\mathbb{R}}\Big(\Phi({\mathbb{U}^\alpha_{\Delta x}}^{(2)}(t^{n+1},x)) - \Phi({\mathbb{U}^\alpha_{\Delta x}}^{(1)}(t^{n+1},x))\Big)\zeta(t^{n+1},x)dx, \tag{5.46}$$

$$\mathcal{E}_S = \sum_{n\ge 0}\int_{\mathbb{R}}\Big(\mathcal{U}(u^\alpha_{\Delta x}(t^{n+1},x)) - \mathcal{U}({u^\alpha_{\Delta x}}^{(3)}(t^{n+1},x))\Big)\zeta(t^{n+1},x)dx. \tag{5.47}$$

First, the dominated convergence theorem readily ensures from the definition of the discrete initial data that

$$\int_{\mathbb{R}_x} \mathcal{U}(u^0_{\Delta x}(x))\zeta(0,x)dx \to \int_{\mathbb{R}_x}\mathcal{U}(u_0(x))\zeta(0,x)dx \quad \text{as } \Delta x \to 0. \tag{5.48}$$

Next we prove that in the limit $\Delta x \to 0$ with $\Delta t/\Delta x$ kept constant

$$\sum_{n\ge 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x} \Phi(u^\alpha_{\Delta x}, v^\alpha_{\Delta x})(t,x)\partial_t\zeta dtdx \to \int_{\mathbb{R}^+_t\times\mathbb{R}_x}\mathcal{U}(u^\alpha)(t,x)\partial_t\zeta dtdx.$$

To this end, we make use of the following triangle inequality:

$$\begin{aligned}&\left|\sum_{n\ge 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\Big(\Phi(u^\alpha_{\Delta x}, v^\alpha_{\Delta x})(t,x) - \mathcal{U}(u^\alpha)(t,x)\Big)\partial_t\zeta dtdx\right| \\ &\quad \le \sum_{n\ge 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\left|\Phi(u^\alpha_{\Delta x}(t,x), v^\alpha_{\Delta x}(t,x)) - \Phi\big(u^\alpha_{\Delta x}(t,x), f(u^\alpha_{\Delta x}(t^n,x))\big)\right||\partial_t\zeta|dtdx \\ &\qquad + \sum_{n\ge 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\Big|\Phi\big(u^\alpha_{\Delta x}(t,x), f(u^\alpha_{\Delta x}(t^n,x))\big) \\ &\qquad\qquad - \Phi\big(u^\alpha_{\Delta x}(t,x), f(u^\alpha_{\Delta x}(t,x))\big)\Big||\partial_t\zeta|dtdx \\ &\qquad + \sum_{n\ge 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\left|\Phi\big(u^\alpha_{\Delta x}(t,x), f(u^\alpha_{\Delta x}(t,x))\big) - \mathcal{U}(u^\alpha)(t,x)\right||\partial_t\zeta|dtdx \\ &\quad := I_1 + I_2 + I_3.\end{aligned} \tag{5.49}$$

Inequality (5.4) together with the sup norm estimate (5.1) yield

(5.50)

$$\begin{aligned}
&\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\left|\Phi(u^\alpha_{\Delta x},v^\alpha_{\Delta x})(t,x)-\Phi\big(u^\alpha_{\Delta x}(t,x),f(u^\alpha_{\Delta x}(t^n,x))\big)\right|\left|\partial_t\zeta dt\right|dx\\
&\quad\leq C\parallel \partial_t\zeta\parallel_{L^\infty(\mathbb{R}_t^+\times\mathbb{R}_x)}\int_{t^n}^{t^{n+1}}\int_{supp(\zeta(t,.))}|v^\alpha_{\Delta x}(t,x)-f(u^\alpha_{\Delta x}(t^n,x))|dxdt\\
&\quad\leq C\int_{t^n}^{t^{n+1}}(s-t^n)ds\,\mathbb{1}_{supp(\zeta(t,.))}\\
&\quad\leq C\Delta t^2\,\mathbb{1}_{supp(\zeta(t,.))},
\end{aligned}$$

where $\mathbb{1}_{supp(\zeta(t,.))}$ denotes the characteristic function of the test function $\zeta(t,.)$ at time t. Hence, if $\Delta t/\Delta x$ is kept constant, we infer

$$I_1\leq C\Delta x, \tag{5.51}$$

where C is independent of Δx. Similarly a combination of (5.3) and (5.1) yields

$$I_2\leq C\Delta x. \tag{5.52}$$

Concerning I_3, since the extracted subsequence $\{u^\alpha_{\Delta x}\}_{\Delta x>0}$ is uniformly bounded in the sup-norm and converges to u^α in $L^\infty\big((0,T),L^1_{loc}(\mathbb{R})\big)$ for all $T>0$ and a.e., the dominated convergence theorem applies and we have

$$\begin{aligned}
&\sum_{n\geq 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\Phi\left(u^\alpha_{\Delta x},f(u^\alpha_{\Delta x})\right)(t,x)\partial_t\zeta dtdx=\int_{\mathbb{R}_t^+\times\mathbb{R}_x}\mathcal{U}(u^\alpha_{\Delta x})(x,t)\partial_t\zeta dtdx\\
&\to\int_{\mathbb{R}_t^+\times\mathbb{R}_x}\mathcal{U}(u^\alpha)(t,x)\partial_t\zeta dtdx
\end{aligned}$$

as $\Delta x\to 0$, thus I_3 vanishes in the reported limit. Exactly the same steps can be applied to show that

$$\sum_{n\geq 0}\int_{t^n}^{t^{n+1}}\int_{\mathbb{R}_x}\Psi\left(u^\alpha_{\Delta x},f(u^\alpha_{\Delta x})\right)(t,x)\partial_x\zeta dtdx\to\int_{\mathbb{R}_t^+\times\mathbb{R}_x}\mathcal{F}(u^\alpha)\partial_x\zeta dtdx,\ \text{as } \Delta x\to 0.$$

Next, we rewrite the averaging error term as follows:

(5.53)

$$\begin{aligned}
&\mathcal{E}_A=\sum_{n\geq 0}\sum_{j\in\mathbb{Z}}(\mathcal{E}_A)^n_j,\\
&(\mathcal{E}_A)^n_j=\int_{x^{n+1}_{j-1/2}}^{x^{n+1}_{j+1/2}}\left(\Phi\left({\mathbb{U}^\alpha_{\Delta x}}^{(2)}(t^{n+1},x)\right)-\Phi\left({\mathbb{U}^\alpha_{\Delta x}}^{(1)}(t^{n+1},x)\right)\right)\zeta(t^{n+1},x)dx.
\end{aligned}$$

Introducing the averaged quantity

$$\zeta^{n+1}_j=\frac{1}{x^{n+1}_{j+1/2}-x^{n+1}_{j-1/2}}\int_{x^{n+1}_{j-1/2}}^{x^{n+1}_{j+1/2}}\zeta(t^{n+1},x)dx, \tag{5.54}$$

with

$$||\zeta(t^{n+1},.)-\zeta^{n+1}_j||_{L^\infty((x^{n+1}_{j-1/2},x^{n+1}_{j+1/2}))}\leq C\Delta x\,\mathbb{1}_{supp(\zeta)}, \tag{5.55}$$

the following identity holds:

$$(\mathcal{E}_A)_j^n = \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \Big(\Phi\Big({\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x)\Big) - \Phi\Big({\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x)\Big)\Big)(\zeta(t^{n+1},x) - \zeta_j^{n+1})dx$$
$$+ \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \Big(\Phi\left({\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x)\right) - \Phi\left({\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x)\right) dx\, \zeta_j^{n+1}$$
$$:= (I_4)_j^{n+1} + (I_5)_j^{n+1}.$$

The convexity of the entropy Φ ensures the following pointwise inequality

(5.56)
$$\Phi\left({\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x)\right) - \Phi\left({\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x)\right)$$
$$-\nabla\Phi\left({\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x)\right)\cdot\left({\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x) - {\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x)\right) \geq 0.$$

Now we make use of the local averaging procedure (4.17) together with the definition (4.21) proposed in the second step, to get the identity

(5.57)
$${\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x) := \mathbb{U}_j^{n+1,(2)}$$
$$= \frac{1}{x_{j+\frac{1}{2}}^{n+1} - x_{j-\frac{1}{2}}^{n+1}} \int_{x_{j-\frac{1}{2}}^{n+1}}^{x_{j+\frac{1}{2}}^{n+1}} {\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x)dx, \quad x \in (x_{j-\frac{1}{2}}^{n+1}, x_{j+\frac{1}{2}}^{n+1}).$$

We thus infer from (5.56) and (5.57) the following bound:

$$(5.58)\quad (I_5)_j^{n+1} \leq \zeta_j^{n+1}\nabla\Phi(\mathbb{U}_j^{n+1,(2)}) \cdot \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \left(\mathbb{U}_j^{n+1,(2)} - {\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x)\right)dx = 0.$$

Then in view of the local error sup-norm estimate stated in (5.55), we deduce

$$(5.59)\qquad (\mathcal{E}_A)_j^n \leq C\Delta x \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} |{\mathbb{U}_{\Delta x}^{\alpha}}^{(2)}(t^{n+1},x) - {\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x)|dx\, \mathbb{1}_{supp(\zeta)}.$$

Invoking the identity (5.57), we get from the uniform BV bounds (5.2):

$$(\mathcal{E}_A)_j^n \leq \frac{C\Delta x\, \mathbb{1}_{supp(\zeta)}}{x_{j+1/2}^{n+1} - x_{j-1/2}^{n+1}} \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} |{\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},x) - {\mathbb{U}_{\Delta x}^{\alpha}}^{(1)}(t^{n+1},y)|dxdy$$
$$\leq TV_{\mathbb{R}}(\mathbb{U}_{\Delta x}^{\alpha}(t^{n+1},.)) \frac{C\Delta x}{x_{j+1/2}^{n+1} - x_{j-1/2}^{n+1}} \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} \int_{x_{j-1/2}^{n+1}}^{x_{j+1/2}^{n+1}} |x-y|dxdy\, \mathbb{1}_{supp(\zeta)}$$
$$\leq C\Delta x^3\, \mathbb{1}_{supp(\zeta)}.$$

Therefore, with $\Delta t/\Delta x$ kept constant, we deduce that the averaging error term is non-positive as Δx goes to zero

$$(5.60)\qquad (\mathcal{E}_A) \leq C\Delta x \sum_{n\geq 0}\sum_{j\in\mathbb{Z}} \mathbb{1}_{supp(\zeta)}\Delta x\Delta t \leq C\Delta x.$$

To conclude, the overall sampling error $\mathcal{E}_S(\Delta x, \alpha, \zeta)$ can be shown to converge to zero as $\Delta x \to 0$ for almost any given sequence $\alpha \in \mathcal{A} = (0,1)^{\mathbb{N}}$, using exactly the same arguments as those developed by Glimm [10] in the convergence analysis of his scheme (see also Serre [25]). With Serre's notations, consider $d\nu(\alpha)$ the measure defined on the Borel sets of the space of sequences $\mathcal{A} = (0,1)^{\mathbb{N}}$, then the estimate

$$\int_{\mathcal{A}} |\mathcal{E}_S(\Delta x, \alpha, \zeta)|^2 \, d\nu(\alpha) \leq C \sup_{t \geq 0} \mathrm{TV}\,(u^{\alpha}_{\Delta x}(t,\cdot))\, \Delta x \leq C \Delta x \tag{5.61}$$

follows, thanks to the property that $u^{\alpha}_{\Delta x}(t,x)$ has uniformly bounded total variation for all $\alpha \in \mathcal{A} = (0,1)^{\mathbb{N}}$. We refer the reader to [25] (Lemma 5.4.2, Chapter 5) for a proof. The proposed estimate actually ensures that for any given test function $\zeta \in \mathcal{C}^1_0(\mathbb{R}^+_t \times \mathbb{R}_x)$, there exists a negligeable set $\mathcal{N}_\zeta \subset \mathcal{A}$ such that for all sequences in $\mathcal{A}/\ \mathcal{N}_\zeta$, the sampling error $\mathcal{E}(\Delta x, \alpha, \zeta)$ goes to zero with Δx. We can therefore conclude that the limit function u^α verifies

$$\int_{\mathbb{R}^+_t \times \mathbb{R}_x} \Big(\mathcal{U}(u^\alpha)\partial_t \zeta + \mathcal{F}(u^\alpha)\partial_x \zeta \Big) dt dx + \int_{\mathbb{R}_x} \mathcal{U}(u_0)\zeta(0,x) dx \geq 0, \tag{5.62}$$

for almost any given sampling sequence $\alpha \in \mathcal{A}$ and for any non-negative test function $\zeta \in C^1_c\left((0,\infty) \times \mathbb{R}_x\right)$. Again, in the case of a genuinely non-linear flux function, the proposed inequality holds for a single strictly convex entropy pair. After Panov [23], it suffices to observe that in addition u^α verifies by construction

$$\int_{\mathbb{R}^+_t \times \mathbb{R}_x} \Big(u^\alpha \partial_t \zeta + f(u^\alpha)\partial_x \zeta \Big) dt dx + \int_{\mathbb{R}_x} u_0 \zeta(0,x) dx = 0, \tag{5.63}$$

namely u^α is a weak solution which satisfies one entropy inequality (5.62): it necessarily coincides with the Kružkov solution. In the situation of a general non-linear flux function, the inequality (5.62) holds for the whole Kružkov family which readily implies that u^α is nothing but the Kružkov solution of the Cauchy problem under consideration. □

6. Numerical examples

In this section we present numerical results to highlight the importance of handling infinitely many entropy pairs in the design of the anti-diffusive law $\Theta(u_L, u_R)$ for a flux function without genuine non-linearity. To this end, we consider the initial value problem

$$\begin{aligned} &\partial_t u + \partial_x \left(\tfrac{u^3}{3} \right) = 0, \quad t > 0, x \in (0,1), \\ &u(0,x) = u_0(x) = \begin{cases} u_L = -1, & x < 0.5, \\ u_R = +1, & x > 0.5, \end{cases} \end{aligned} \tag{6.1}$$

with Neumann boundary conditions. The exact solution of this Riemann problem is a compound wave made of a shock attached to a rarefaction wave, as depicted in the figures displayed hereafter. The initial data in (6.1) is chosen so that the entropy jump for the quadratic entropy pair is zero

$$-\sigma(u_L, u_R)(\frac{u_R^2}{2} - \frac{u_L^2}{2}) + (\frac{u_R^4}{4} - \frac{u_L^4}{4}) = 0, \quad \sigma(u_L, u_R) = \frac{1}{3}. \tag{6.2}$$

Hence choosing the anti-diffusive law (3.24) designed for genuinely non-linear flux functions comes with $\Gamma(u_L, u_R) = 0$ in (3.25) so that the optimal value $\Theta(u_L, u_R)$ in (3.24) boils down to 1. With such a law any of the two methods in sections 4.1 and 4.2 capture a weak solution made of a single discontinuity propagating with speed $\sigma(u_L, u_R) = 1/3$. This weak solution is entropy violating. It is therefore of central importance to promote the anti-diffusive law (3.62) to enforce the validity of all the Kružkov entropy inequalities. Numerical results displayed below assess these issues.

The solution of the IBVP (6.1) is approximated using the Jin-Xin method with and without defect measure corrections to illustrate their relative performance. With defect measure, we implemented both Algorithms 4.1 and 4.2. The anti-diffusive law is first set to the optimal law (3.62) especially designed for general non-linear flux function. It is then set to (3.24) as a comparison for our numerical purposes. In the calculations, we use the low variance van der Corput sequence $\alpha \equiv \{\alpha_n\}_{n\geq 0}$ (see [15] for instance) defined by

$$\alpha_n = \sum_{k=0}^{m} i_k 2^{-(k+1)}, \quad \text{with } n = \sum_{k=0}^{m} i_k 2^k, \tag{6.3}$$

where the i_k represents the binary expansion of the integer $n = 1, 2, \ldots$. The first few elements of this sequence are

$$\begin{aligned} &\alpha_1 = 0.5, && \alpha_2 = 0.25, && \alpha_3 = 0.75, && \alpha_4 = 0.125, \\ &\alpha_5 = 0.625, && \alpha_6 = 0.375, && \alpha_7 = 0.875, && \alpha_8 = 0.0625. \end{aligned} \tag{6.4}$$

The number of points in space is taken to be 250 and the CFL condition is set at the value of 0.45. Exact and discrete solutions for the Jin-Xin method without defect measure corrections are compared in Figure 1. Corresponding results for the Jin-Xin method with defect measure corrections based on the optimal law (3.62) are displayed in Figures 2 and 3. Observe the fairly good agreement achieved with the exact solution. Results obtained for the optimal law (3.24) are plotted in Figure 4. As expected, the method captures a wrong weak solution.

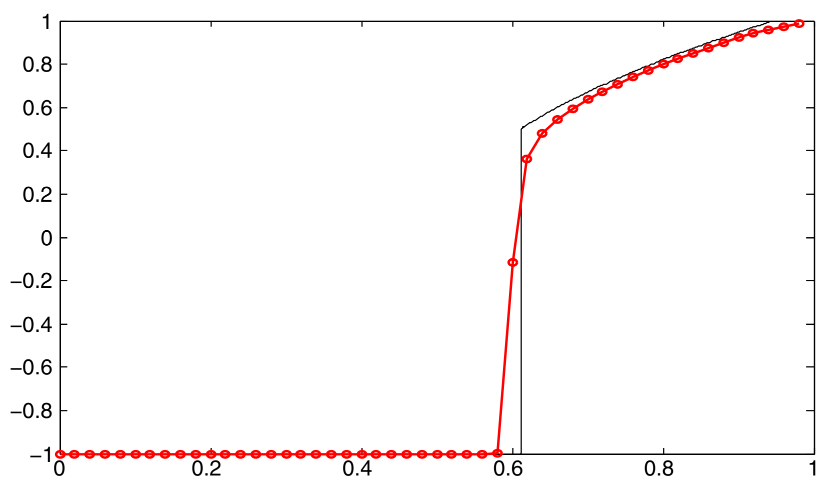

FIGURE 1. Example (6.1). Jin-Xin method without defect measure corrections. Solid line: exact solution; circles: numerical solution.

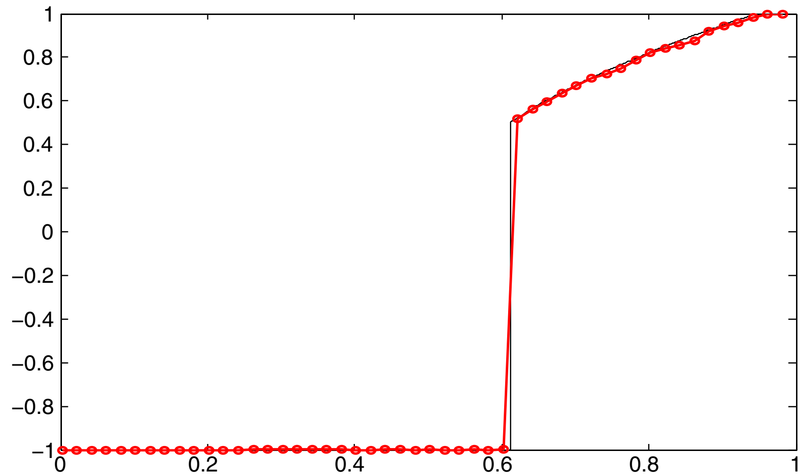

FIGURE 2. Example (6.1). Algorithm 4.1: Jin-Xin method with defect measure corrections based on the anti-diffusive law (3.62) for a general non-linear flux. Solid line: exact solution; circles: numerical solution.

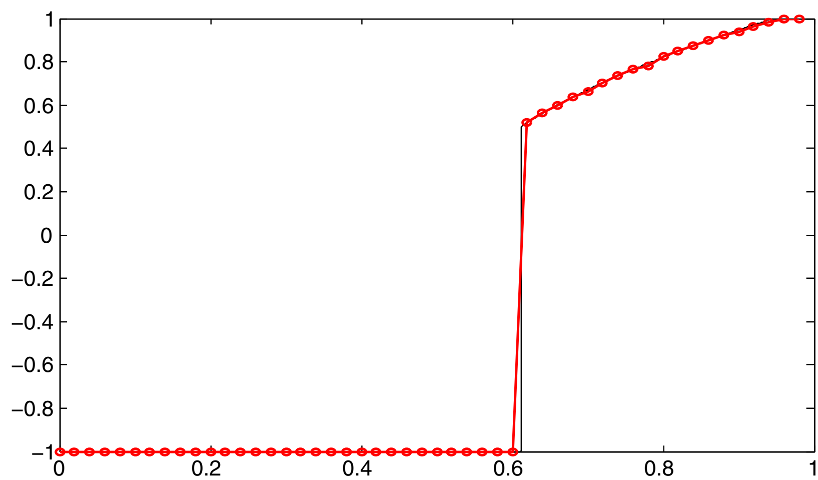

FIGURE 3. Example (6.1). Algorithm 4.2: Jin-Xin method with defect measure corrections based on the anti-diffusive law (3.62) for a general non-linear flux. Solid line: exact solution; circles: numerical solution.

In the end, we give another example constituting two shocks meeting each other and finally merging into one. The problem is stated as follows:

$$\begin{aligned} &\partial_t u + \partial_x \left(\frac{u^2}{2}\right) = 0, \quad t > 0, x \in (0,1), \\ &u(0,x) = u_0(x) = \begin{cases} u_L = -1, & x < 0.25, \\ u_R = +1, & 0.25 \le x < 0.75, \\ u_R = -1, & x \ge 0.75, \end{cases} \end{aligned} \tag{6.5}$$

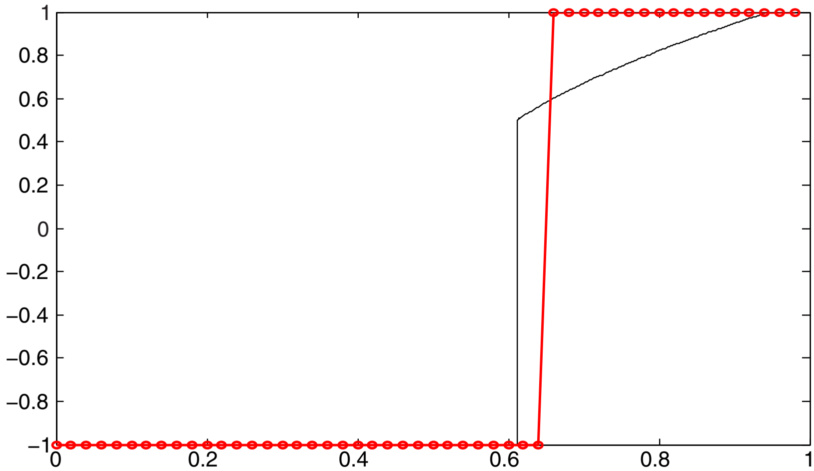

FIGURE 4. Example (6.1). Jin-Xin method with defect measure corrections based on the anti-diffusive law (3.24) designed for a genuinely non-linear flux. Solid line: exact solution; circles: numerical solution.

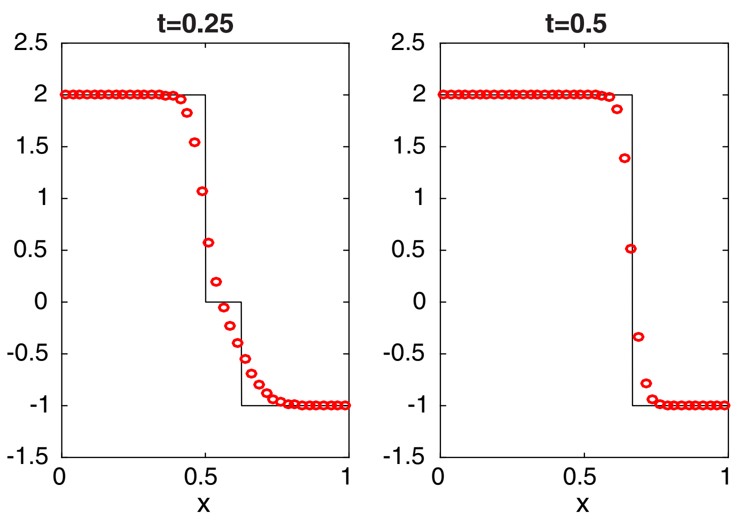

FIGURE 5. Example (6.5). Jin-Xin method without defect measure corrections. Solid line: exact solution; circles: numerical solution.

and the Neumann boundary condition is used. We choose mesh size as $\Delta x = 0.025$ and set CFL number to be 0.2. The results are displayed in Figures 5, 6, and 7 and the output times are $t = 0.25$ and $t = 0.5$. In all figures, the black curve is the exact solution and red circle denotes the numerical solution. Here both Algorithms 4.1 and 4.2 capture shock very well whereas the Jix-Xin method without defect measure produces smeared results.

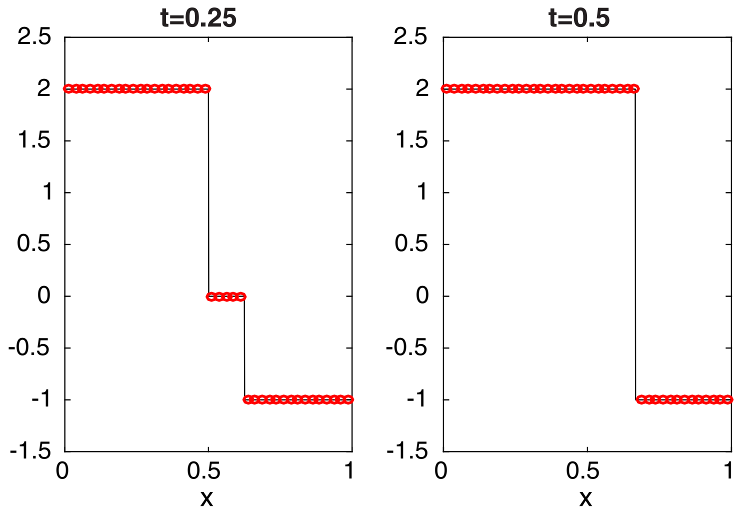

FIGURE 6. Example (6.5). Algorithm 4.1: Jin-Xin method with defect measure corrections based on the anti-diffusive law (3.62) for a general non-linear flux. Solid line: exact solution; circles: numerical solution.

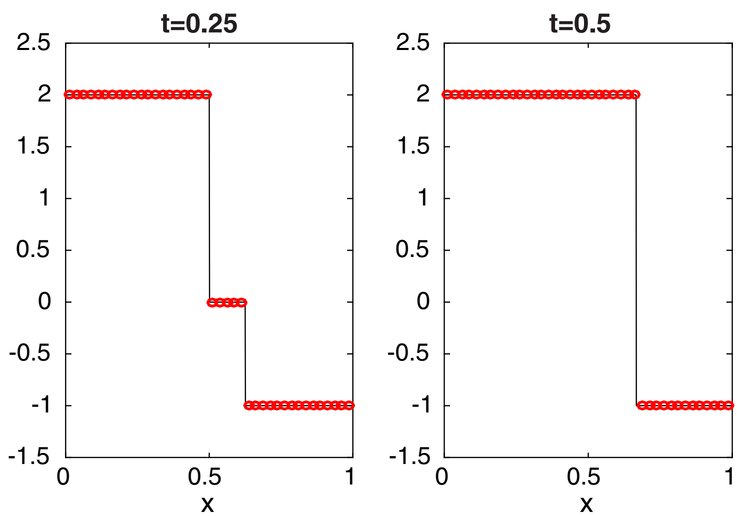

FIGURE 7. Example (6.5). Algorithm 4.2: Jin-Xin method with defect measure corrections based on the anti-diffusive law (3.62) for a general non-linear flux. Solid line: exact solution; circles: numerical solution.

REFERENCES

[1] P. Baiti, A. Bressan, and H. K. Jenssen, *Instability of travelling wave profiles for the Lax-Friedrichs scheme*, Discrete Contin. Dyn. Syst. **13** (2005), no. 4, 877–899, DOI 10.3934/dcds.2005.13.877. MR2166709

[2] A. Bressan, H. K. Jenssen, and P. Baiti, *An instability of the Godunov scheme*, Comm. Pure Appl. Math. **59** (2006), no. 11, 1604–1638, DOI 10.1002/cpa.20141. MR2254446

[3] W. Bao and S. Jin, *The random projection method for hyperbolic conservation laws with stiff reaction terms*, J. Comput. Phys. **163** (2000), no. 1, 216–248, DOI 10.1006/jcph.2000.6572. MR1777727

[4] J. Boris and D. Book, *Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works*, J. Comput. Phys. **11** (1973), no. 1, 38–69.

[5] C. Chalons and F. Coquel, *Modified Suliciu relaxation system and exact resolution of isolated shock waves*, Math. Models Methods Appl. Sci. **24** (2014), no. 5, 937–971, DOI 10.1142/S0218202513500723. MR3187187

[6] C. Chalons, F. Coquel, P. Engel, and C. Rohde, *Fast relaxation solvers for hyperbolic-elliptic phase transition problems*, SIAM J. Sci. Comput. **34** (2012), no. 3, A1753–A1776, DOI 10.1137/110848815. MR2970272

[7] G. Q. Chen, C. D. Levermore, and T.-P. Liu, *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Comm. Pure Appl. Math. **47** (1994), no. 6, 787–830, DOI 10.1002/cpa.3160470602. MR1280989

[8] J. Cheng and C.-W. Shu, *Second order symmetry-preserving conservative Lagrangian scheme for compressible Euler equations in two-dimensional cylindrical coordinates*, J. Comput. Phys. **272** (2014), 245–265, DOI 10.1016/j.jcp.2014.04.031. MR3212271

[9] P. Colella, A. Majda, and V. Roytburd, *Theoretical and numerical structure for reacting shock waves*, SIAM J. Sci. Statist. Comput. **7** (1986), no. 4, 1059–1080, DOI 10.1137/0907073. MR857783

[10] J. Glimm, *Solutions in the large for nonlinear hyperbolic systems of equations*, Comm. Pure Appl. Math. **18** (1965), 697–715, DOI 10.1002/cpa.3160180408. MR0194770

[11] E. Godlewski and P.-A. Raviart, *Hyperbolic systems of conservation laws*, Mathématiques and Applications (Paris) [Mathematics and Applications], vol. 3/4, Ellipses, Paris, 1991. MR1304494

[12] A. Harten, *High resolution schemes for hyperbolic conservation laws*, J. Comput. Phys. **49** (1983), no. 3, 357–393, DOI 10.1016/0021-9991(83)90136-5. MR701178

[13] A. Harten, *ENO schemes with subcell resolution*, J. Comput. Phys. **83** (1989), no. 1, 148–184, DOI 10.1016/0021-9991(89)90226-X. MR1010163

[14] A. Harten and J. M. Hyman, *Self-adjusting grid methods for one-dimensional hyperbolic conservation laws*, J. Comput. Phys. **50** (1983), no. 2, 235–269, DOI 10.1016/0021-9991(83)90066-9. MR707200

[15] A. Harten and P. D. Lax, *A random choice finite difference scheme for hyperbolic conservation laws*, SIAM J. Numer. Anal. **18** (1981), no. 2, 289–315, DOI 10.1137/0718021. MR612144

[16] A. Harten, P. D. Lax, and B. van Leer, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev. **25** (1983), no. 1, 35–61, DOI 10.1137/1025002. MR693713

[17] S. Jin and J.-G. Liu, *The effects of numerical viscosities. I. Slowly moving shocks*, J. Comput. Phys. **126** (1996), no. 2, 373–389, DOI 10.1006/jcph.1996.0144. MR1404378

[18] S. Jin and Z. P. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Comm. Pure Appl. Math. **48** (1995), no. 3, 235–276, DOI 10.1002/cpa.3160480303. MR1322811

[19] S. N. Kružkov, *First order quasilinear equations with several independent variables.* (Russian), Mat. Sb. (N.S.) **81 (123)** (1970), 228–255. MR0267257

[20] P. G. LeFloch and S. Mishra, *Numerical methods with controlled dissipation for small-scale dependent shocks*, Acta Numer. **23** (2014), 743–816, DOI 10.1017/S0962492914000099. MR3202244

[21] R. Natalini, *Convergence to equilibrium for the relaxation approximations of conservation laws*, Comm. Pure Appl. Math. **49** (1996), no. 8, 795–823, DOI 10.1002/(SICI)1097-0312(199608)49:8⟨795::AID-CPA2⟩3.0.CO;2-3. MR1391756

[22] W. F. Noh, *Errors for calculations of strong shocks using an artificial viscosity and an artificial heat flux*, J. Comp. Phys. **72** (1987), 78–120.

[23] E. Yu. Panov, *Uniqueness of the solution of the Cauchy problem for a first-order quasilinear equation with an admissible strictly convex entropy* (Russian, with Russian summary), Mat. Zametki **55** (1994), no. 5, 116–129, 159, DOI 10.1007/BF02110380; English transl., Math. Notes **55** (1994), no. 5-6, 517–525. MR1296003

[24] J. J. Quirk, *A contribution to the great Riemann solver debate*, Internat. J. Numer. Methods Fluids **18** (1994), no. 6, 555–574, DOI 10.1002/fld.1650180603. MR1270805

[25] D. Serre, *Systems of conservation laws. 2*, Cambridge University Press, Cambridge, 2000. Geometric structures, oscillations, and initial-boundary value problems; Translated from the 1996 French original by I. N. Sneddon. MR1775057

[26] C.-W. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997), Lecture Notes in Math., vol. 1697, Springer, Berlin, 1998, pp. 325–432, DOI 10.1007/BFb0096355. MR1728856

[27] C.-W. Shu, *A brief survey on discontinuous Galerkin methods in computational fluid dynamics*, Advances in Mechanics **43** (2013), 541–554.

[28] M. Slemrod, *Admissibility criteria for propagating phase boundaries in a van der Waals fluid*, Arch. Rational Mech. Anal. **81** (1983), no. 4, 301–315, DOI 10.1007/BF00250857. MR683192

[29] B. van Leer, *Towards the ultimate conservative difference scheme. v. a second-order sequel to godunov's method*, J. Comput. Phys. **32** (1979), no. 1, 101–136.

CNRS and Centre de Mathématiques Appliquées UMR 7641, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France

E-mail address: `frederic.coquel@cmap.polytechnique.fr`

Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, Wisconsin 53706

E-mail address: `jin@math.wisc.edu`

Department of Physics and Department of Mathematics, Duke University, Durham, North Carolina 27708

E-mail address: `Jian-Guo.Liu@duke.edu`

Department of Mathematics and Computational Data-Enabled Science and Engineering Program, State University of New York at Buffalo, 244 Mathematics Building, Buffalo, New York 14260

E-mail address: `lwang46@buffalo.edu`