Found Comput Math (2017) 17:763–827
DOI 10.1007/s10208-015-9299-z

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Construction of Approximate Entropy Measure-Valued Solutions for Hyperbolic Systems of Conservation Laws

Ulrik S. Fjordholm[1] · **Roger Käppeli**[2] · **Siddhartha Mishra**[2] · **Eitan Tadmor**[3]

Received: 5 February 2014 / Revised: 23 June 2015 / Accepted: 3 November 2015 /
Published online: 21 December 2015

Abstract Entropy solutions have been widely accepted as the suitable solution framework for systems of conservation laws in several space dimensions. However, recent results in De Lellis and Székelyhidi Jr (Ann Math 170(3):1417–1436, 2009) and Chiodaroli et al. (2013) have demonstrated that entropy solutions may not be unique. In this paper, we present numerical evidence that state-of-the-art numerical schemes *need not* converge to an entropy solution of systems of conservation laws as the mesh is refined. Combining these two facts, we argue that entropy solutions may not be suitable as a solution framework for systems of conservation laws, particularly in several space dimensions. We advocate *entropy measure-valued solutions*, first proposed by DiPerna, as the appropriate solution paradigm for systems of conservation laws. To this end, we present a detailed numerical procedure which constructs stable

Communicated by Wolfgang Dahmen.

✉ Eitan Tadmor
tadmor@cscamm.umd.edu

Ulrik S. Fjordholm
ulrik.fjordholm@math.ntnu.no

Roger Käppeli
roger.kaeppeli@sam.math.ethz.ch

Siddhartha Mishra
smishra@sam.math.ethz.ch

1 Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway

2 Seminar for Applied Mathematics, ETH Zürich, HG J 48, Rämistrasse 101, Zurich, Switzerland

3 Center for Scientific Computation and Mathematical Modeling (CSCAMM), Department of Mathematics, Institute for Physical Science and Technology (IPST), University of Maryland, College Park, MD 20742-4015, USA

approximations to entropy measure-valued solutions, and provide sufficient conditions that guarantee that these approximations converge to an entropy measure-valued solution as the mesh is refined, thus providing a viable numerical framework for systems of conservation laws in several space dimensions. A large number of numerical experiments that illustrate the proposed paradigm are presented and are utilized to examine several interesting properties of the computed entropy measure-valued solutions.

Keywords Hyperbolic conservation laws · Uniqueness · Stability · Entropy condition · Measure-valued solutions · Atomic initial data · Random field · Weak BV estimate · Weak* convergence

Mathematics Subject Classification 65M06 · 35L65 · 35R06

"There is no theory for the initial value problem for compressible flows in two space dimensions once shocks show up, much less in three space dimensions. This is a scientific scandal and a challenge."
P. D. Lax, 2007 Gibbs Lecture [48].

Contents

1 Introduction

A large number of problems in physics and engineering are modeled by *systems of conservation laws*

$$\partial_t u + \nabla_x \cdot f(u) = 0 \tag{1a}$$

$$u(x,0) = u_0(x). \tag{1b}$$

Here, the unknown $u = u(x,t) : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^N$ is the vector of *conserved variables* and $f = (f^1, \ldots, f^d) : \mathbb{R}^N \to \mathbb{R}^{N\times d}$ is the *flux function*. We denote $\mathbb{R}_+ := [0, \infty)$.

The system (1a) is *hyperbolic* if the flux Jacobian $\partial_u(f \cdot n)$ has real eigenvalues for all $n \in \mathbb{R}^d$ with $|n| = 1$. Examples of hyperbolic systems of conservation laws include the shallow water equations of oceanography, the Euler equations of gas dynamics, the magnetohydrodynamics (MHD) equations of plasma physics, the equations of nonlinear elastodynamics and the Einstein equations of general relativity. We refer to [16,37] for more theory on hyperbolic conservation laws.

1.1 Mathematical Framework

It is well known that solutions of the Cauchy problem (1) can develop discontinuities such as shock waves in finite time, even when the initial data is smooth. Hence, solutions of hyperbolic systems of conservation laws (1) are sought in the weak (distributional) sense.

Definition 1 A function $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$ is a *weak solution* of (1) if it satisfies (1) in the sense of distributions:

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}^d} \partial_t\varphi(x,t)u(x,t) + \nabla_x\varphi(x,t)\cdot f(u(x,t))\,\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d}\varphi(x,0)u_0(x)\,\mathrm{d}x = 0 \tag{2}$$

for all test functions $\varphi \in C_c^1(\mathbb{R}^d \times \mathbb{R}_+)$.

Weak solutions are in general not unique: Infinitely many weak solutions may exist after the formation of discontinuities. Thus, to obtain uniqueness, additional admissibility criteria have to be imposed. These admissibility criteria take the form of entropy conditions, which are formulated in terms of entropy pairs.

Definition 2 A pair of functions (η, q) with $\eta : \mathbb{R}^N \to \mathbb{R}$, $q : \mathbb{R}^N \to \mathbb{R}^d$ is called an *entropy pair* if η is convex and q satisfies the compatibility condition $q' = \eta' \cdot f'$.

Definition 3 A weak solution u of (1) is an *entropy solution* if the entropy inequality

$$\partial_t\eta(u) + \nabla_x \cdot q(u) \leqslant 0 \quad \text{in } \mathcal{D}'\big(\mathbb{R}^d \times \mathbb{R}_+\big)$$

is satisfied for all entropy pairs (η, q), that is, if

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}^d} \partial_t\varphi(x,t)\eta(u(x,t))+\nabla_x\varphi(x,t)\cdot q(u(x,t))\,\mathrm{d}x\mathrm{d}t+\int_{\mathbb{R}^d}\varphi(x,0)\eta(u_0(x))\,\mathrm{d}x\geqslant 0 \tag{3}$$

for all nonnegative test functions $0\leqslant\varphi\in C_c^1(\mathbb{R}^d\times\mathbb{R}_+)$.

For the special case of scalar conservation laws ($N=1$), every convex function η gives rise to an entropy pair by letting $q(u):=\int^u\eta'(\xi)f'(\xi)\mathrm{d}\xi$. This rich family of entropy pairs was used by Kruzkhov [45] to obtain existence, uniqueness and stability of solutions for scalar conservation laws.

Corresponding (global) well-posedness results for systems of conservation laws are much harder to obtain. Lax [47] showed existence and stability of entropy solutions for one-dimensional systems of conservation laws for the special case of Riemann initial data. Existence results for the Cauchy problem for one-dimensional systems were obtained by Glimm [35] using the random choice method and by Bianchini and Bressan [5] with the vanishing viscosity method. Uniqueness and stability results for one-dimensional systems were shown by Bressan et al. [9]. All of these results rely on an assumption that the initial data is "sufficiently small," i.e., lies sufficiently close to a constant state.

On the other hand, *no global existence and uniqueness (stability) results are currently available* for a generic system of conservation laws in several space dimensions. In fact, recent results (see [17–19] and references therein) provide counterexamples which illustrate that entropy solutions for multi-dimensional systems of conservation laws *are not necessarily unique*. These results raise serious questions whether the notion of entropy solutions is too restricted to serve as the standard solution framework for systems of conservation laws. It can be argued that one needs to impose even further admissibility criteria, in addition to the entropy inequality (3), to single out a solution among the infinitely many solutions constructed in [17–19]. One possible approach in determining these selection criteria is to employ suitable numerical schemes and observe which, if any, of the entropy solutions are approximated by these schemes.

1.2 Numerical Schemes

Numerical schemes have played a leading role in the study of systems of conservation laws, and a wide variety of numerical methods for approximating (1) are currently available. These include the very popular and highly successful numerical framework of finite volume and finite difference schemes, based on approximate Riemann solvers or on Riemann-solver-free central differencing (see [10,13,37,50]) which utilize TVD [38], ENO [39] or WENO [42] non-oscillatory reconstruction techniques and strong stability preserving (SSP) Runge–Kutta time integrators [34]. Another popular alternative is the discontinuous Galerkin finite element method [14].

The primary goal in the analysis of numerical schemes approximating (1) is proving convergence to an entropy solution as the mesh is refined. This issue has been addressed in the special case of (first-order) monotone schemes for *scalar* conservation laws (see [15] for the one-dimensional case and [12] for multiple dimensions) using the TVD property. Corresponding convergence results for (formally) arbitrarily high-order accurate finite difference schemes for scalar conservation laws were obtained recently in [28], see also [27]. Convergence results for (arbitrarily high order) space time DG discretization for scalar conservation laws were obtained in [41] and for the spectral viscosity method in [62].

The question of convergence of numerical schemes for systems of conservation laws is significantly more difficult. Currently, there are *no* rigorous proofs of convergence for any kind of finite volume (difference) and finite element methods to the entropy solutions of a generic system of conservation laws, even in one space dimension. Convergence aside, even the stability of numerical approximations to systems of conservation laws is mostly open. The only notion of numerical stability for systems of conservation laws that has been analyzed rigorously so far is that of entropy stability—the design of numerical approximations that satisfy a discrete version of the entropy inequality. Such schemes have been devised in [27,40,60,61]. However, entropy stability may not suffice to ensure the convergence of approximate solutions.

1.3 Two Numerical Experiments

Given the lack of rigorous stability and convergence results for systems of conservation laws, it has become customary in the field to rely on numerical benchmark tests to demonstrate the convergence of the scheme empirically. One such benchmark test case is the radial Sod shock tube [50].

1.3.1 Sod Shock Tube

In this test, we consider the compressible Euler equations of gas dynamics in two space dimensions (see Sect. 6) as a prototypical hyperbolic system of conservation laws. The initial data for the two-dimensional version of the well-known Sod shock tube problem is given by

$$u_0(x) = \begin{cases} u_L & \text{if } |x| \leqslant r_0 \\ u_R & \text{if } |x| > r_0, \end{cases} \tag{4}$$

with $\rho_L = p_L = 3$, $\rho_R = p_R = 1$, $w^1 = w^2 = 0$. The computational domain is $[-0.5, 0.5]^2$, with $r_0 = 0.15$, and we use periodic boundary conditions.

To begin with, we consider a perturbed version of the Sod shock tube by setting the initial data

$$u_0^\varepsilon(x) = u_0(x) + \varepsilon X(x), \tag{5}$$

where $\varepsilon > 0$ is a small amplitude of the perturbation $X(\cdot)$ associated with the following state variables—ρ, p and $w = (w^1, w^2)^\top$,

$$X_\rho = X_p = 0, \quad X_{w^1}(x) = \sin(2\pi x_1), \quad X_{w^2}(x) = \sin(2\pi x_2). \tag{6}$$

First we set $\varepsilon = 0.01$ and compute the approximate solutions of the two-dimensional Euler equations (31) with the second-order TeCNO2 finite difference scheme of [27]. In Fig. 1, we present the computed densities at time $t = 0.24$ for three different mesh resolutions. The figure clearly indicates convergence as the mesh is refined. To further quantify this convergence, we compute the difference in the approximate solutions on two successive mesh resolutions:

$$E^{\Delta x} = \left\| u^{\Delta x} - u^{\Delta x/2} \right\|_{L^1([-0.5,0.5]^2)}, \tag{7}$$

and plot the results for density in Fig. 2a. The results clearly indicate that the numerical approximations form a Cauchy sequence in L^1 and hence converge. The same numerical experiment was performed with a different scheme: a second-order high-resolution scheme based on an HLLC solver using the MC limiter, implemented in the FISH code [44]. Similar convergence results were obtained (omitted here for brevity).

Next, we investigate numerically the issue of stability of this system with respect to perturbations in the initial data. To this end, we use exactly the same set up as the previous numerical experiment but let the perturbation amplitude $\varepsilon \to 0$ in (5) and plot in Fig. 2b the error in computed density (at a fixed mesh resolution of 1024^2 points) for successively lower values of ε. The reference solution is computed with the finest mesh resolution of 1024^2 using the unperturbed initial data (4). The results clearly show convergence to the unperturbed solution in the $\varepsilon \to 0$ limit.

The above numerical example suggests convergence of the approximate numerical solutions to an entropy solution, at least for some benchmark test cases. The computed solutions were observed to be stable with respect to perturbations of the initial data. In the literature, it is not uncommon to extrapolate from benchmark test cases like the Sod shock tube and expect that the numerical approximations converge as the mesh is refined for all possible sets of flow configurations.

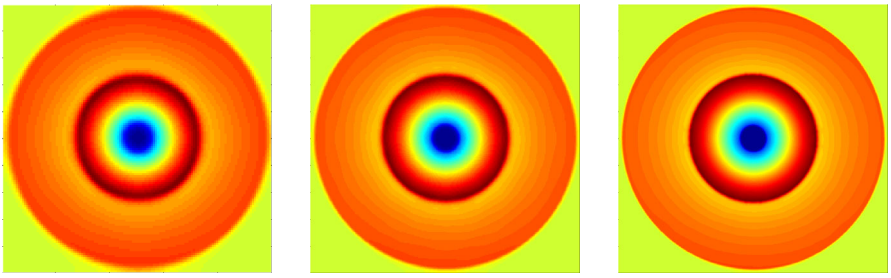


Fig. 1 Density for the Sod shock tube problem, computed with TECNO2 finite difference scheme of [27], with initial data (5) at time $t = 0.24$. *Left* to *right* $\Delta x = 1/128, 1/256, 1/512$

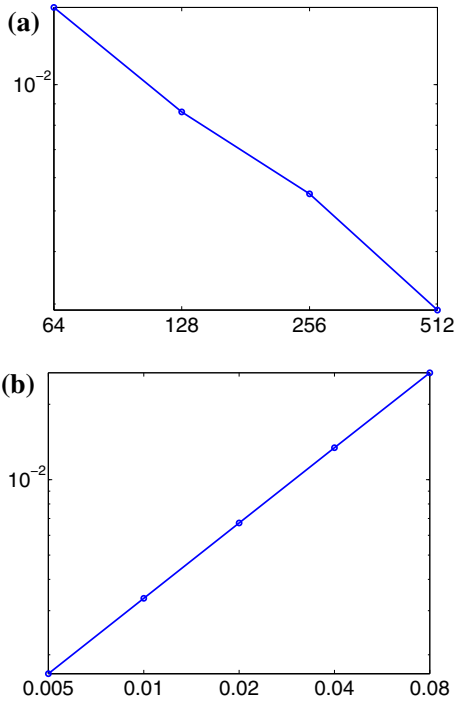


Fig. 2 L^1 differences in density ρ at time $t = 0.24$ for the Sod shock tube problem with initial data (5). **a** L^1 Cauchy rates (7) (y-axis) in the density at time $t = 0.24$ versus number of gridpoints (x-axis). **b** L^1 error with respect to the unperturbed solution (4) (y-axis) versus the perturbation parameter ε (x-axis)

1.3.2 Kelvin–Helmholtz Problem

We question the universality of the above observed empirical convergence and stability results by considering the following set of initial data for the two-dimensional Euler equations (see Sect. 6):

$$u_0(x) = \begin{cases} u_L & \text{if } 0.25 < x_2 < 0.75 \\ u_R & \text{if } x_2 \leqslant 0.25 \text{ or } x_2 \geqslant 0.75, \end{cases} \tag{8}$$

with $\rho_L = 2$, $\rho_R = 1$, $w_L^1 = -0.5$, $w_R^1 = 0.5$, $w_L^2 = w_R^2 = 0$ and $p_L = p_R = 2.5$. It is readily seen that this is a steady state, i.e., that $u(x, t) \equiv u_0(x)$ is an entropy solution.

Next, we add the same perturbation (5) to the initial data (8) and compute approximate solutions in the computational domain $[0, 1]^2$ with periodic boundary conditions, for different $\Delta x > 0$. A series of approximate solutions using the TeCNO2 scheme of [27] and perturbation amplitude $\varepsilon = 0.01$ are shown in Fig. 3. The results show that there is no sign of any convergence as the mesh is refined. As a matter of fact, structures at smaller and smaller scales are formed with mesh refinement. This lack of convergence is quantified by plotting the differences between successive mesh levels (7) for the density in Fig. 4a. The results show that as the mesh is refined, the approximate solutions do *not* seem to form a Cauchy sequence in L^1 (at least for the mesh

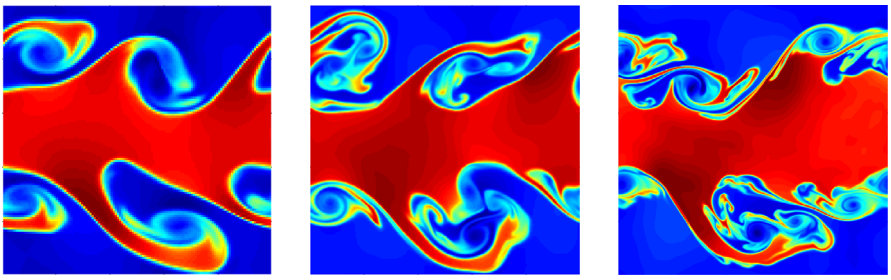


Fig. 3 Density for the Kelvin–Helmholtz problem (8) with perturbation (5) and perturbation parameter $\varepsilon = 0.01$. *Left* to *right* $\Delta x = 1/128, 1/256, 1/512$, at time $t = 1$

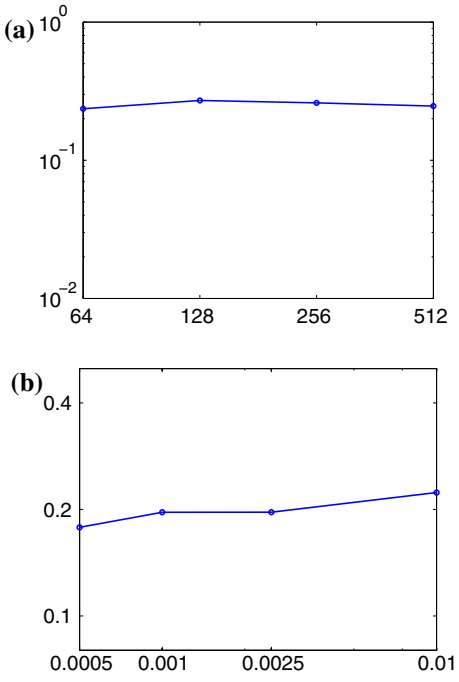


Fig. 4 L^1 differences in density ρ at time $t = 2$ for the Kelvin–Helmholtz problem (8). **a** L^1 Cauchy rates (7) (y-axis) versus number of gridpoints (x-axis) for the perturbed problem (5), (8) with $\varepsilon = 0.01$. **b** L^1 error with respect to the steady-state solution (8) of the unperturbed Kelvin–Helmholtz problem (y-axis) versus perturbation parameter ε, at a fixed mesh with 1024^2 points

resolutions that have been tested) and hence may not converge. The results presented in Figs. 3 and 4a are computed with the TeCNO scheme of [27]. Very similar results were also obtained with the FISH code [44] and the ALSVID finite volume code [31]. Furthermore, convergence in even weaker $W^{-1,p}$, $1 < p \leqslant \infty$, norms was also not observed. Thus, one cannot deduce convergence of even bulk properties of the flow, such as the average domain temperature, in this particular case.

Finally, we check stability of the numerical solutions as the perturbation parameter $\varepsilon \to 0$. We compute numerical approximations at a fixed fine grid resolution of 1024^2 points with successively lower values of ε. These results are compared with the steady-state solution (8) and are presented in Fig. 4b. The L^1 difference results clearly show that there is no convergence to the steady-state solution (8) as $\varepsilon \to 0$.

1.4 A Different Notion of Solutions

The above experiment clearly demonstrates that in general, a whole host[1] of state-of-the-art numerical schemes do not seem to converge (even for very fine mesh resolutions) to an entropy solution for multi-dimensional systems of conservation laws. In fact, structures at smaller and smaller scales are formed as the mesh is refined. This fact does not imply that the numerical approximations are at fault (given that all the tested schemes, based on different design philosophies, behaved in the same manner), but rather, that the notion of entropy solutions does not adequately describe the complex flow phenomena that are modeled by systems of conservation laws such as the compressible Euler equations.

When combined with the recent results on the non-uniqueness of entropy solutions of systems of conservation laws ([17,18] and references therein), our numerical evidence strongly suggests, in more than one way, that entropy solutions need not be an appropriate solution framework for systems of conservation laws. In particular, entropy solutions may not suffice to characterize the limits of numerical approximations to conservation laws in a stable manner.

Based on the fact that oscillations persist on finer and finer scales in numerical approximations of (1) shown in Fig. 3, we focus on the alternative concept of *entropy measure-valued solutions*, introduced by DiPerna in [22], see also [23]. In this framework, solutions of the system of conservation laws (1) are no longer integrable functions, but parameterized probability measures, or *Young measures*, which are able to represent the limit behavior of sequences of oscillatory functions. This solution concept was further based on the work of Tartar [63] on characterizing the weak limits of bounded sequences of functions. More recently, Glimm and co-workers ([11,51] and references therein) have also hypothesized that entropy measure-valued solutions are the appropriate notion of solutions for hyperbolic conservation laws, particularly in several space dimensions.

1.5 Aims and Scope of the Current Paper

In the current paper:

- We consider entropy measure-valued solutions for the Cauchy problem (1), in the sense of DiPerna [22]. We study the existence and stability of the entropy measure-valued solutions.
- The main aim of the current paper is to approximate entropy measure-valued solutions numerically. To this end, we propose an algorithm based on the realization of Young measures as the law of random fields and approximate the solution random fields with suitable finite difference (volume) numerical schemes. We propose a set of sufficient conditions that a scheme has to satisfy in order to

[1] We have tested at least three types of schemes, TeCNO scheme of [27], the high-resolution HLLC scheme of [44] and the finite volume scheme of [31], and obtained similar non-convergence and instability results
Footnote 1 continued
as presented above. We strongly suspect that any numerical method will not converge or be stable with respect to perturbations in the initial data for this particular example.

converge to an entropy measure-valued solution as the mesh is refined. Examples of such convergent schemes are also provided. This provides a viable and rigorous numerical framework for multi-dimensional systems of conservation laws, within the framework of entropy measure-valued solutions.
– We present a large number of numerical experiments to validate the proposed theory. The numerical approximations are also employed to study the stability as well as other interesting properties of entropy measure-valued solutions.

The rest of this paper is organized as follows: In Sect. 2, we provide a short but self-contained description of Young measures (see also Appendix 1) and then define entropy measure-valued solutions for a generalized Cauchy problem, corresponding to the system of conservation law (1). The well posedness of the entropy measure-valued solutions is discussed in Sect. 3. In Sect. 4, we discuss finite difference schemes approximating (1) and propose abstract criteria that these schemes have to satisfy in order to converge to entropy measure-valued solutions. Two schemes satisfying the abstract convergence framework are presented in Sect. 5. In Sect. 6, we present numerical experiments that illustrate the convergence properties of the schemes and discuss the stability and related properties of entropy measure-valued solutions.

2 Young Measures and Entropy Measure-Valued Solutions

A *Young measure* on a set $D \subset \mathbb{R}^k$ (in our setting, $D = \mathbb{R}^d \times \mathbb{R}_+$ will represent space–time) is a function ν which assigns to every point $y \in D$ a probability measure $\nu_y \in \mathcal{P}(\mathbb{R}^N)$ on the phase space $\mathbb{R}^N$. The set of all Young measures from D to $\mathbb{R}^N$ is denoted by $\mathbf{Y}(D, \mathbb{R}^N)$. We can compose a Young measure with a continuous function g by defining $\langle \nu_y, g \rangle := \int_{\mathbb{R}^N} g(\xi) \mathrm{d}\nu_y(\xi)$, the expectation of g with respect to the probability measure ν_y. Note that this defines a real-valued function of $y \in D$.

Every measurable function $u : D \to \mathbb{R}^N$ gives rise to a Young measure by letting

$$\nu_y := \delta_{u(y)},$$

where δ_ξ is the Dirac measure centered at $\xi \in \mathbb{R}^N$. Such Young measures are called *atomic*.

If $\nu^1, \nu^2, \ldots$ is a sequence of Young measures, then there are two notions of convergence. We say that ν^n converge *weak** to a Young measure ν (written $\nu^n \rightharpoonup \nu$) if $\langle \nu^n, g \rangle \overset{*}{\rightharpoonup} \langle \nu, g \rangle$ in $L^\infty(D)$ for all $g \in C_0(\mathbb{R}^N)$, that is, if

$$\int_D \varphi(z) \left\langle \nu_z^n, g \right\rangle \, \mathrm{d}z \to \int_D \varphi(z) \, \langle \nu_z, g \rangle \, \mathrm{d}z \quad \forall \, \varphi \in L^1(D). \tag{9}$$

By the *fundamental theorem of Young measures* (see Theorem 13), *any* suitably bounded sequence of Young measures has a weak* convergent subsequence.

We say that the sequence $\{\nu^n\}$ converges *strongly* to ν (written $\nu^n \to \nu$) if

$$\left\| W_p(\nu^n, \nu) \right\|_{L^p(D)} \to 0 \tag{10}$$

for some $p \in [1, \infty)$, where W_p is the *p-Wasserstein distance*

$$W_p(\mu, \rho) := \inf \left\{ \int_{\mathbb{R}^N \times \mathbb{R}^N} |\xi - \zeta|^p \, \mathrm{d}\pi(\xi, \zeta) \ : \ \pi \in \Pi(\mu, \rho) \right\}^{1/p}$$

which metrizes the topology of weak convergence on the set $\mathcal{P}^p(\mathbb{R}^N) := \{\mu \in \mathcal{P}(\mathbb{R}^N) : \langle \mu, |\xi|^p \rangle < \infty\}$. Here, $\Pi(\mu, \rho)$ is the set of probability measures on $\mathbb{R}^N \times \mathbb{R}^N$ with marginals $\mu, \rho \in \mathcal{P}^p(\mathbb{R}^N)$ (see also Appendix A.1.3).

We refer to Appendix 1 for a more rigorous and detailed introduction to Young measures.

2.1 The Measure-Valued (MV) Cauchy Problem

As mentioned in the introduction, we will seek a more general (weaker) notion of solutions to the Cauchy problem for a system of conservation laws (1) by requiring that the solutions be Young measures, instead of integrable functions. Equipped with the notation of the previous section, we propose the following generalized Cauchy problem (corresponding to the system (1)): find a $\nu \in \mathbf{Y}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$ such that

$$\begin{aligned} \partial_t \langle \nu, \mathrm{id} \rangle + \nabla_x \cdot \langle \nu, f \rangle &= 0 \\ \nu_{(x,0)} &= \sigma_x, \end{aligned} \tag{11}$$

where $\sigma \in \mathbf{Y}(\mathbb{R}^d, \mathbb{R}^N)$ is the *initial measure-valued data* and $\mathrm{id}(\xi) = \xi$ is the identity function on $\mathbb{R}^N$. The above MV Cauchy problem is interpreted as follows.

Definition 4 (DiPerna [22]) A Young measure $\nu \in \mathbf{Y}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$ is a *measure-valued (MV) solution* of (11) if (11) holds in the sense of distributions, i.e.,

$$\begin{aligned} &\int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \partial_t \varphi(x,t) \left\langle \nu_{(x,t)}, \mathrm{id} \right\rangle + \nabla_x \varphi(x,t) \cdot \left\langle \nu_{(x,t)}, f \right\rangle \, \mathrm{d}x\mathrm{d}t \\ &\quad + \int_{\mathbb{R}^d} \varphi(x,0) \, \langle \sigma_x, \mathrm{id} \rangle \, \mathrm{d}x = 0 \end{aligned} \tag{12}$$

for all test functions $\varphi \in C_c^1(\mathbb{R}^d \times \mathbb{R}_+)$.

Definition 5 (DiPerna [22]) A Young measure $\nu \in \mathbf{Y}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$ is an *entropy measure-valued (EMV) solution* of (11) if in addition to being a measure-valued solution (satisfying (12)), it also satisfies

$$\partial_t \langle \nu, \eta \rangle + \nabla_x \cdot \langle \nu, q \rangle \leqslant 0 \quad \text{in } \mathcal{D}'\big(\mathbb{R}^d \times \mathbb{R}_+\big) \tag{13}$$

for every entropy pair (η, q), that is, if

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}^d} \partial_t\varphi(x,t)\left\langle \nu_{(x,t)}, \eta\right\rangle + \nabla_x\varphi(x,t)\cdot\left\langle \nu_{(x,t)}, q\right\rangle \,\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d}\varphi(x,0)\left\langle \sigma_x, \eta\right\rangle \,\mathrm{d}x \geqslant 0 \tag{14}$$

for all nonnegative test functions $0 \leqslant \varphi \in C_c^1(\mathbb{R}^d \times \mathbb{R}_+)$.

Remark 1 The formulations (12) and (14) impose the initial data σ in a *very weak* manner. The weak formulation (12) requires, roughly speaking, that $\lim_{t\to 0}\left\langle \nu_{(x,t)}, \mathrm{id}\right\rangle = \langle \sigma_x, \mathrm{id}\rangle$, i.e., that the *barycenters* (or mean) of $\nu_{(x,0)}$ and σ_x should coincide. The inequality (14) implies that $\limsup_{t\to 0}\left\langle \nu_{(x,t)}, \eta\right\rangle \leqslant \langle \sigma_x, \eta\rangle$ (in Theorem 2, we require a slightly stronger form of this inequality). The requirement that the (strictly) η-weighted barycenters of two measures should "coincide" (up to the entropic inequality) will imply that the measures themselves coincide *only* if σ_x is a Dirac mass. Correspondingly, our condition for initial data implies equality at $t = 0$ *only* when the initial data is atomic. This is precisely the setting which we choose to focus on in the present paper. In a forthcoming paper [25], we consider the question of uniqueness when the initial data is non-atomic, and how to interpret the initial condition in this more complex setting.

We denote by $\mathcal{E}(\sigma)$ the set of all entropy MV solutions of the MV Cauchy problem (11) with initial MV data σ. It is readily seen that every entropy solution u of (1) gives rise to an EMV solution of (11) with $\sigma = \delta_{u_0}$, by defining $\nu_{(x,t)} := \delta_{u(x,t)}$, the atomic Young measure concentrated at u. Thus, the set $\mathcal{E}(\sigma)$ is at least as large as the set of entropy solutions of (1) whenever σ is atomic, $\sigma = \delta_{u_0}$.

Remark 2 Although our focus in the current paper will be on the specific case of atomic initial data, we still consider the more general setting of the MV Cauchy problem (11) as it enables us to formulate numerical approximations in a unified manner.

In practice, the initial data u_0 in (1a) is obtained from a measurement or observation process. Since measurements (observations) are intrinsically uncertain, it is customary to model this initial uncertainty statistically by considering the initial data u_0 as a random field. Given the fact that the law of a random field is a Young measure, we can also model this initial uncertainty with non-atomic initial measures in the measure-valued (MV) Cauchy problem (11). Thus, our formulation suffices to include various formalisms for uncertainty quantification of conservation laws, i.e., the determination of solution uncertainty given uncertain initial data. See [52–54] and references therein for an extensive discussion on uncertainty quantification for conservation laws.

3 Well Posedness of EMV Solutions

The questions of existence, uniqueness and stability of EMV solutions of (11) are of fundamental significance. We start with a discussion of the scalar case.

3.1 Scalar Conservation Laws

The question of existence of EMV solutions for scalar conservation laws was considered by DiPerna [22]. We slightly generalize his result for a non-atomic initial data as follows.

Theorem 1 *Consider the MV Cauchy problem* (11) *for a scalar conservation law. If the initial data* σ *is uniformly bounded (see* Appendix A.2.2*), then there exists an EMV solution of* (11).

Proof By Proposition 2 in Appendix A.3.2, there exists a probability space $(\Omega, \mathcal{F}, P)$ and a random field $u_0 : \Omega \times \mathbb{R}^d \to \mathbb{R}$ with law σ. By the uniform boundedness of σ, we have $\|u_0\|_{L^\infty(\Omega\times\mathbb{R}^d)} < \infty$.

For each $\omega \in \Omega$, let $u(\omega; x, t)$ be the entropy solution of (1) with initial data $u_0(\omega)$, and define ν as the law of u. Then for every entropy pair (η, q) and every test function $0 \leqslant \varphi \in C_c^1(\mathbb{R}^d \times \mathbb{R}_+)$, we have by Fubini's theorem and the entropy stability of $u(\omega)$ for each ω,

$$\begin{aligned}
&\int_{\mathbb{R}_+}\int_{\mathbb{R}^d} \partial_t\varphi(x,t)\left\langle \nu_{(x,t)}, \eta\right\rangle + \nabla_x\varphi(x,t)\cdot\left\langle \nu_{(x,t)}, q\right\rangle \,\mathrm{d}x\mathrm{d}t\\
&\quad= \int_{\mathbb{R}_+}\int_{\mathbb{R}^d} \partial_t\varphi(x,t)\int_\Omega \eta(u(\omega;x,t))\,\mathrm{d}P(\omega)\\
&\qquad+ \nabla_x\varphi(x,t)\cdot\int_\Omega q(u(\omega;x,t))\,\mathrm{d}P(\omega)\mathrm{d}x\mathrm{d}t\\
&\quad= \int_\Omega\int_{\mathbb{R}_+}\int_{\mathbb{R}^d} \partial_t\varphi(x,t)\eta(u(\omega;x,t)) + \nabla_x\varphi(x,t)\cdot q(u(\omega;x,t))\,\mathrm{d}x\mathrm{d}t\mathrm{d}P(\omega)\\
&\quad\geqslant -\int_\Omega\int_{\mathbb{R}^d} \varphi(x,0)\eta(u_0(\omega;x))\,\mathrm{d}x\mathrm{d}P(\omega)\\
&\quad= -\int_{\mathbb{R}^d} \varphi(x,0)\left\langle \sigma_x, \eta\right\rangle \,\mathrm{d}x.
\end{aligned}$$

This proves the entropy inequality (14). □

Although EMV solutions exist for scalar conservation laws with non-atomic measure-valued initial data, they may not be unique. Here is a simple counter-example (see also Schochet [58]).

Example 1 Consider Burgers' equation

$$\partial_t u + \partial_x\left(\frac{u^2}{2}\right) = 0.$$

Denote by λ the Lebesgue measure on $\mathbb{R}$ and by λ_A the restriction of λ to a subset $A \subset \mathbb{R}$, i.e., $\lambda_A(B) = \lambda(A \cap B)$. We define $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$ (the Borel σ-algebra on $[0, 1]$) and $P = \lambda_{[0,1]}$. Let u_0 and $\tilde{u}_0$ be the random fields

$$u_0(\omega; x) := \begin{cases} 1+\omega & \text{for } x < 0 \\ \omega & \text{for } x > 0, \end{cases}$$

$$\tilde{u}_0(\omega; x) := \begin{cases} 1+\omega & \text{for } x < 0 \\ 1-\omega & \text{for } x > 0, \end{cases} \qquad \omega \in [0,1],\ x \in \mathbb{R}.$$

It is readily checked that the law of both u_0 and $\tilde{u}_0$ in $(\Omega, \mathcal{F}, P)$ equals

$$\sigma_x = \begin{cases} \lambda_{[1,2]} & \text{for } x < 0 \\ \lambda_{[0,1]} & \text{for } x > 0. \end{cases}$$

Note that both of the random fields u_0 and $\tilde{u}_0$ model the Burgers' equation with *uncertainty in initial shock location* which is widely considered in the UQ literature, see [52] and references therein. Although their laws are the same—i.e., that the initial Young measure is the same in both cases—the resulting two-point correlations (particularly for points left and right of the origin) differ.

The entropy solutions $u(\omega)$ and $\tilde{u}(\omega)$ of the Riemann problems with initial data $u_0(\omega)$ and $\tilde{u}_0(\omega)$ are given by

$$u(\omega; x, t) = \begin{cases} 1+\omega & \text{if } x/t < 1/2 + \omega \\ \omega & \text{if } x/t > 1/2 + \omega, \end{cases} \qquad \tilde{u}(\omega; x, t) = \begin{cases} 1+\omega & \text{if } x/t < 1 \\ 1-\omega & \text{if } x/t > 1. \end{cases}$$

To compute the law ν of u, we rewrite u as

$$u(\omega; x, t) = \begin{cases} 1+\omega & \text{if } x/t - 1/2 < \omega \\ \omega & \text{if } x/t - 1/2 > \omega. \end{cases}$$

Hence, if $x/t - 1/2 < 0$, then $\nu_{(x,t)} = \lambda_{[1,2]}$, whereas if $x/t - 1/2 > 1$, then $\nu_{(x,t)} = \lambda_{[0,1]}$. When $0 \leqslant x/t - 1/2 \leqslant 1$, we have for every $g \in C_0(\mathbb{R}^N)$

$$\begin{aligned} \langle \nu_{(x,t)}, g \rangle &= \int_0^1 g(u(\omega; x, t))\, \mathrm{d}\omega = \int_{x/t-1/2}^1 g(1+\omega)\, \mathrm{d}\omega + \int_0^{x/t-1/2} g(\omega)\, \mathrm{d}\omega \\ &= \int_{x/t+1/2}^2 g(\omega)\, \mathrm{d}\omega + \int_0^{x/t-1/2} g(\omega)\, \mathrm{d}\omega \\ &= \int_{\mathbb{R}} g(\omega)\, \mathrm{d}\lambda_{[x/t+1/2,2]}(\omega) + \int_{\mathbb{R}} g(\omega)\, \mathrm{d}\lambda_{[0,x/t-1/2]}(\omega). \end{aligned}$$

After a similar calculation for $\tilde{\nu}$, we find that

$$\nu_{(x,t)} = \begin{cases} \lambda_{[1,2]} & \text{if } x/t < 1/2 \\ \lambda_{[x/t+1/2,2]} + \lambda_{[0,x/t-1/2]} & \text{if } 1/2 < x/t < 3/2 \\ \lambda_{[0,1]} & \text{if } 3/2 < x/t, \end{cases} \qquad \tilde{\nu}_{(x,t)} = \begin{cases} \lambda_{[1,2]} & \text{if } x/t < 1 \\ \lambda_{[0,1]} & \text{if } x/t > 1. \end{cases}$$

Note that in fact $\nu_{(x,t)}$ and $\tilde{\nu}_{(x,t)}$ converges to σ_x *strongly* as $t \to 0$ for all $x \neq 0$. Thus, ν and $\tilde{\nu}$ are EMV solutions with the same initial MV data σ, but do not coincide. □

The above example clearly illustrates that the MV Cauchy problem (11) may not have unique solutions, even for the scalar case, when the initial data is a non-atomic Young measure. Hence, it raises serious questions whether the notion of an entropy measure-valued solution is useful. However, the following result shows that when restricting attention to the relevant class of *atomic initial data*, EMV solutions of the *scalar* MV Cauchy problem (11) are stable.

Theorem 2 *Consider the scalar case $N = 1$. Let $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$ and let $\sigma \in \mathbf{Y}(\mathbb{R}^d)$ be uniformly bounded. Let $u \in L^1 \cap L^\infty(\mathbb{R}^d \times \mathbb{R}_+)$ be the entropy solution of the scalar conservation law (1) with initial data u_0. Let ν be **any** EMV solution of (11) which satisfies*

$$\limsup_{T\to 0} \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \left\langle \nu_{(x,t)}, |u(x,t) - \xi| \right\rangle \, dxdt \leqslant \int_{\mathbb{R}^d} \langle \sigma_x, |u_0(x) - \xi| \rangle \, dx. \quad (15)$$

Then for all $t > 0$,

$$\int_{\mathbb{R}^d} \left\langle \nu_{(x,t)}, |u(x,t) - \xi| \right\rangle \, dx \leqslant \int_{\mathbb{R}^d} \langle \sigma_x, |u_0(x) - \xi| \rangle \, dx,$$

or equivalently,

$$\left\| W_1\big(\nu_{(\cdot,t)}, \delta_{u(\cdot,t)}\big) \right\|_{L^1(\mathbb{R}^d)} \leqslant \left\| W_1\big(\sigma, \delta_{u_0}\big) \right\|_{L^1(\mathbb{R}^d)}.$$

In particular, if $\sigma = \delta_{u_0}$ then $\nu = \delta_u$.

Proof We follow DiPerna [22] who proved the uniqueness of scalar MV solutions subject to atomic initial data. Here, we quantify stability in terms of the W_1-metric, which is related to the $L^1(x, v)$-stability of *kinetic* solutions associated with (1), see [57].

For $\xi \in \mathbb{R}$, let $(\eta(\xi, u), q(\xi, u))$ be the Kruzkov entropy pair, defined as

$$\eta(\xi, u) := |\xi - u|, \quad q(\xi, u) := \operatorname{sgn}(\xi - u)(f(\xi) - f(u)), \quad u, \xi \in \mathbb{R}.$$

By [22, Theorem 4.1] we know that for any entropy solution u of (1) and any entropy MV solution ν of (11), we have

$$\partial_t \left\langle \nu_{(x,t)}, \eta\big(\xi, u(x,t)\big) \right\rangle + \nabla_x \cdot \left\langle \nu_{(x,t)}, q\big(\xi, u(x,t)\big) \right\rangle \leqslant 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (0,\infty)),$$

that is,

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \left(\partial_t \varphi(x,t) \int_{\mathbb{R}^N} \eta\big(\xi, u(x,t)\big) \, \mathrm{d}\nu_{(x,t)}(\xi) \right.$$
$$\left. + \nabla_x \varphi(x,t) \cdot \int_{\mathbb{R}^N} q\big(\xi, u(x,t)\big) \, \mathrm{d}\nu_{(x,t)}(\xi) \right) \mathrm{d}x\mathrm{d}t \geqslant 0$$

for all test functions $0 \leqslant \varphi \in C_c^1\big(\mathbb{R}^d \times (0,\infty)\big)$. Setting $\varphi(x,t) = \theta(t)$ for a $\theta \in C_c^\infty\big((0,\infty)\big)$, we get

$$\int_{\mathbb{R}_+} \theta'(t) V(t)\, \mathrm{d}t \geqslant 0, \quad V(t) := \int_{\mathbb{R}^d} \big\langle \nu_{(x,t)}, |\xi - u(x,t)| \big\rangle \, \mathrm{d}x.$$

Letting θ be a smooth approximation of the indicator function on an interval $[0, t_0]$, we find in light of (15) that $V(t_0) \leqslant \int_{\mathbb{R}^d} \langle \sigma_x, |u_0(x) - \xi| \rangle \, \mathrm{d}x$ for almost every $t_0 > 0$. □

3.2 Systems of Conservation Laws

It is clear from the above discussion that non-atomic initial data might lead to multiple EMV solutions, see also the discussions in Remark 2.3. However, the scalar results also suggest some possible stability with respect to perturbations of *atomic* initial data. Based on these considerations, we propose the following (weaker) notion of stability.

Terminology 3 *The MV Cauchy problem* (11) *is* MV stable *if the following property holds.*

For every $u_0 \in L^\infty(\mathbb{R}^d, \mathbb{R}^N)$ *and* $\sigma \in \mathbf{Y}(\mathbb{R}^d, \mathbb{R}^N)$ *such that*

$$\mathscr{D}\left(\delta_{u_0}, \sigma\right) \ll 1,$$

there exists an EMV solution $\nu \in \mathcal{E}(\delta_{u_0})$ *such that*

$$\mathscr{D}\left(\nu, \nu^\sigma\right) \ll 1$$

for every EMV solution $\nu^\sigma \in \mathcal{E}(\sigma)$ *(or a subset thereof).*

(Recall that $\mathcal{E}(\sigma)$ denotes the set of all entropy MV solutions to the MV Cauchy problem (11).) We have intentionally left out several details in the above definition: the admissible set of initial data; the subset of $\mathcal{E}(\cdot)$ for which the MV Cauchy problem is stable; and the distance $\mathscr{D}$ on the set of Young measures. Still, the concept of MV stability carries one of the main messages in this paper: Despite the well-documented instability of entropic weak solutions, as shown for example in the introduction and in Sect. 6, one could still hope for a stable solution of systems of conservation laws, when it is interpreted as a measure-valued solution, subject to atomic initial data.

Carrying out the full scope of this paradigm for general systems of conservation laws is currently beyond reach. Instead, we examine the question of whether EMV solutions of selected systems of conservation laws are stable or not with the aid of numerical experiments reported in Sect. 6. As for the analytical aspects, we recall that in the scalar case, measure-valued perturbations of atomic initial data are stable (Theorem 2). In the following theorem, we prove the MV stability in the case of systems, provided we further limit ourselves to MV perturbations of *classical solutions* of (11). The proof, along the lines of [20, Theorem 2.2], implies weak–strong uniqueness, as in [8]. In

particular, the theorem provides consistency of EMV solutions with classical solutions of (1), as long as the latter exists.

Theorem 4 *Assume that there exists a classical solution $u \in W^{1,\infty}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$ of (1) with initial data u_0, both taking values in a compact set $K \subset \mathbb{R}^N$. Let ν be an EMV solution of (11) such that the supports of both ν and its initial MV data σ are contained in K. Assume that η is uniformly convex on K. Then for all $t > 0$, there exists a constant C depending on u, such that*

$$\int_{\mathbb{R}^d} \left\langle \nu_{(x,t)}, |u(x,t) - \xi|^2 \right\rangle \, dx \leqslant e^{Ct} \int_{\mathbb{R}^d} \left\langle \sigma_x, |u_0(x) - \xi|^2 \right\rangle \, dx,$$

or equivalently,

$$\left\| W_2\big(\nu_{(\cdot,t)}, \delta_{u(\cdot,t)}\big) \right\|_{L^2(\mathbb{R}^d)} \leqslant e^{Ct} \left\| W_2\big(\sigma, \delta_{u_0}\big) \right\|_{L^2(\mathbb{R}^d)}.$$

In particular, if $\sigma = \delta_{u_0}$ then $\nu = \delta_u$, and so any (classical, weak or measure-valued) solution must coincide with u.

Proof Denote $\overline{u} := \langle \nu, \mathrm{id} \rangle$ and $\overline{u}_0 := \langle \sigma, \mathrm{id} \rangle$. Define the *entropy variables* $v = v(x,t) := \eta'(u(x,t))$ and denote $v_0 := v(x,0) = \eta'(u_0)$. It is readily verified that $v_t = -(f^i)'(u)\partial_i v$ (where $\partial_i = \frac{\partial}{\partial_{x_i}}$). Here and in the remainder of the proof, we use the Einstein summation convention.

Subtracting (12) from (2) and putting $\varphi(x,t) = v(x,t)\theta(t)$ for some $\theta \in C_c^1(\mathbb{R}_+)$ gives

$$\begin{aligned}
0 &= \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} (\overline{u} - u) \cdot \big(v_t \theta + v\theta'\big) + \Big(\big\langle \nu, f^i \big\rangle - f^i(u)\Big) \cdot \partial_i v \theta \, \mathrm{d}x\mathrm{d}t \\
&\quad + \int_{\mathbb{R}^d} (\overline{u}_0 - u_0) \cdot v_0 \theta(0) \, \mathrm{d}x \\
&= \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} (\overline{u} - u) \cdot v\theta' + \left(\underbrace{\big\langle \nu, f^i \big\rangle - f^i(u) - (f^i)'(u)(\overline{u} - u)}_{=: Z^i} \right) \cdot \partial_i v \theta \, \mathrm{d}x\mathrm{d}t \\
&\quad + \int_{\mathbb{R}^d} (\overline{u}_0 - u_0) \cdot v_0 \theta(0) \, \mathrm{d}x
\end{aligned}$$

Next, note that since u is a classical solution, the entropy inequality (3) is in fact an equality. Hence, subtracting (14) from (3) and putting $\varphi(x,t) = \theta(t)$ gives

$$0 \leqslant \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \big(\langle \nu, \eta \rangle - \eta(u)\big)\theta' \, \mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d} \big(\langle \sigma, \eta \rangle - \eta(u_0)\big)\theta(0) \, \mathrm{d}x.$$

Subtracting these two expressions thus gives

$$0 \leqslant \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \hat{\eta}\theta' - Z^i \cdot \partial_i v \theta \, \mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d} \hat{\eta}_0 \theta(0) \, \mathrm{d}x. \tag{16}$$

where

$$\hat{\eta} := \langle v, \eta \rangle - \eta(u) - (\overline{u} - u) \cdot v, \quad \hat{\eta}_0 := \langle \sigma, \eta \rangle - \eta(u_0) - (\overline{u}_0 - u_0) \cdot v_0.$$

Let $\delta > 0$, and let $t > 0$ be a Lebesgue point for the function $s \mapsto \int_{\mathbb{R}} \hat{\eta}(x, s) \, \mathrm{d}x$. We define

$$\theta(s) := \begin{cases} 1 & s < t \\ 1 - \frac{s-t}{\delta} & t \leqslant s < t + \delta \\ 0 & t + \delta \leqslant s. \end{cases}$$

Taking the limit $\delta \to 0$ in (16) then gives

$$\int_{\mathbb{R}^d} \hat{\eta}(t, x) \, \mathrm{d}x \leqslant - \int_0^t \int_{\mathbb{R}^d} Z^i \cdot \partial_i v \, \mathrm{d}x\mathrm{d}s + \int_{\mathbb{R}^d} \hat{\eta}_0 \, \mathrm{d}x.$$

Since $\nu_{(x,s)}$ is a probability distribution, it follows from the uniform convexity of η that

$$\begin{aligned} \hat{\eta} &= \int_K \eta(\xi) - \eta(u) - \eta'(u) \cdot (\xi - u) \, \mathrm{d}\nu_{(x,s)} \\ &\geqslant c \int_K |u - \xi|^2 \, \mathrm{d}\nu_{(x,s)} = c \left\langle \nu_{(x,s)}, |u - \xi|^2 \right\rangle. \end{aligned}$$

Similarly, by the L^∞ bound on both u and $\partial_i v$, we have

$$\hat{\eta}_0 \leqslant C \left\langle \sigma, |u_0 - \xi|^2 \right\rangle \quad \text{and} \quad |Z^i \cdot \partial_i v| \leqslant C \left\langle \nu, |u - \xi|^2 \right\rangle.$$

Hence,

$$\begin{aligned} &\int_{\mathbb{R}^d} \left\langle \nu_{(x,t)}, |u - \xi|^2 \right\rangle \mathrm{d}x \\ &\leqslant C \int_0^t \int_{\mathbb{R}} \left\langle \nu_{(x,s)}, |u - \xi|^2 \right\rangle \mathrm{d}x\mathrm{d}s + C \int_{\mathbb{R}^d} \left\langle \sigma_x, |u_0 - \xi|^2 \right\rangle \mathrm{d}x. \end{aligned}$$

By the integral form of Grönwall's lemma, we obtain the desired result. □

Remark 3 In addition to proving consistency of entropy measure-valued solutions with classical solutions (when they exist), the above theorem also provides local (in time) uniqueness of MV solutions in the following sense. Let $u_0 \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^N)$ be the initial data in (1), then by standard results [16], we have local (in time) existence

of a unique classical solution $u \in W^{1,\infty}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$. By the above theorem, δ_u is also the unique EMV solution of the MV Cauchy problem (11) with initial data δ_{u_0}. However, uniqueness can break down once this MV solution develops singularities.

4 Construction of Approximate EMV Solutions

Although existence results for specific systems of conservation laws such as polyconvex elastodynamics [20], two-phase flows [32,33] and transport equations [11] are available, there exists no global existence result for a generic system of conservation laws. We pursue a different approach by constructing approximate EMV solutions and proving their convergence. A procedure for constructing approximate EMVs is outlined in the present section. It provides a constructive proof of existence of EMV solutions for a generic system of conservation laws, and it is implemented in the numerical simulations reported in Sect. 6.

4.1 Numerical Approximation of EMV Solutions

The construction of approximate EMV solutions consists of several ingredients. It begins with a proper choice of a numerical scheme for approximating the system of conservation laws (1).

4.1.1 Numerical Schemes for One- and Multi-dimensional Conservation Laws

For simplicity, we begin with the description of a numerical scheme for a one-dimensional system of conservation laws, (1) with $d = 1$. We discretize our computational domain into cells $\mathcal{C}_i := [x_{i-1/2}, x_{i+1/2})$ with mesh size $\Delta x = x_{i+1/2} - x_{i-1/2}$ and midpoints

$$x_i := \frac{x_{i-1/2} + x_{i+1/2}}{2}.$$

Note that we consider a uniform mesh size Δx only for the sake of simplicity of the exposition. Next, we discretize the one-dimensional system, $\partial_t u + \partial_x f(u) = 0$, with the following semi-discrete finite difference scheme for $u_i^{\Delta x}(t) \equiv u^{\Delta x}(x_i, t)$ (cf. [37,50]):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} u_i^{\Delta x}(t) + \frac{1}{\Delta x}\left(F^{\Delta x}_{i+1/2}(t) - F^{\Delta x}_{i-1/2}(t)\right) = 0 \quad t > 0,\ i \in \mathbb{Z} \\ u_i^{\Delta x}(0) = u_0^{\Delta x}(x_i) \quad i \in \mathbb{Z}. \end{aligned} \tag{17a}$$

Here, $u_0^{\Delta x}$ is an approximation to the initial data u_0. Henceforth, the dependence of u and F on Δx will be suppressed for notational convenience. The *numerical flux function* $F_{i+1/2}(t)$ is a function depending on $u(x_j, t)$ for $j = i - p + 1, \ldots, i + p$ for some $p \in \mathbb{N}$. It is assumed to be consistent with f and locally Lipschitz continuous, i.e., for every compact $K \subset \mathbb{R}^N$ there is a $C > 0$ such that

$$|F_{i+1/2}(t) - f(u_i(t))| \leqslant C \sum_{j=i-p+1}^{i+p} |u_j - u_i|$$

whenever $u(x_j, t) \in K$ for $j = i - p + 1, \ldots, i + p$.

The semi-discrete scheme (17a) needs to be integrated in time to define a fully discrete numerical approximation. Again for simplicity, we will use an exact time integration, resulting in

$$u_i^{\Delta x}(t + \Delta t) = u_i^{\Delta x}(t) - \frac{1}{\Delta x} \int_t^{t+\Delta t} \left(F_{i+1/2}(\tau) - F_{i-1/2}(\tau)\right) \mathrm{d}\tau. \tag{17b}$$

The function $t \mapsto u(x_i, t)$ is then Lipschitz, that is,

$$\left|u^{\Delta x}(x_i, t) - u^{\Delta x}(x_i, s)\right| \leqslant \frac{C}{\Delta x}|t - s| \quad \forall\, i \in \mathbb{Z},\ t, s \in [0, T].$$

In particular, for all $\Delta x > 0$ and $i \in \mathbb{N}$, the function $t \mapsto u(x_i, t)$ is differentiable almost everywhere. We denote the evolution operator associated with the one-dimensional scheme (17) with mesh size Δx by $\mathcal{S}^{\Delta x}$, so that $u^{\Delta x} = \mathcal{S}^{\Delta x} u_0$.

A similar framework applies to systems of conservation laws in several space dimensions. To simplify the notation, we restrict ourselves to the two-dimensional case (with the usual relabeling $(x_1, x_2) \mapsto (x, y)$), $\partial_t u + \partial_x f^x(u) + \partial_y f^y(u) = 0$.

We discretize our two-dimensional computational domain with into cells with mesh size $\Delta := (\Delta x_1, \Delta x_2)$: With the usual relabeling $(\Delta x_1, \Delta x_2) \mapsto (\Delta x, \Delta y)$), these two-dimensional cells $\mathcal{C}_{i,j} := [x_{i-1/2}, x_{i+1/2}) \times [y_{j-1/2}, y_{j+1/2})$ are assumed to a have a fixed mesh ratio, $\Delta x = x_{i+1/2} - x_{i-1/2}$ and $\Delta y = y_{j+1/2} - y_{j-1/2}$, such that $\Delta y = c\Delta x$ for some constant c. Let

$$(x_i, y_j) = \left(\frac{x_{i-1/2} + x_{i+1/2}}{2}, \frac{y_{j-1/2} + y_{j+1/2}}{2}\right)$$

denote the mid-cells. We end up with the following semi-discrete finite difference scheme for $u_{ij}^{\Delta} = u^{\Delta}(x_i, y_j, t)$ (cf. [37,50]):

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t} u_{ij}^{\Delta}(t) + \frac{1}{\Delta x}\left(F^{x,\Delta x}_{i+1/2,j}(t) - F^{x,\Delta x}_{i-1/2,j}(t)\right) \\ &\quad + \frac{1}{\Delta y}\left(F^{y,\Delta y}_{i,j+1/2}(t) - F^{y,\Delta y}_{i,j-1/2}(t)\right) = 0, \quad t > 0, \\ &u_{ij}^{\Delta}(0) = u_0^{\Delta}(x_i, y_j) \quad i \in \mathbb{Z}. \end{aligned} \tag{18a}$$

Here, $u_0^{\Delta} \approx u_0$ is the approximate initial data and $F^{x,\Delta x}_{i+1/2,j}$, $F^{y,\Delta y}_{i,j+1/2}$ are the locally Lipschitz numerical flux functions which are assumed to be consistent with the flux function $f = (f^x, f^y)$. We integrate the semi-discrete scheme (18a) exactly in time to obtain

$$u_{ij}^{\Delta}(t+\Delta t) = u_{ij}^{\Delta}(t) - \frac{1}{\Delta x}\int_t^{t+\Delta t}\left(F_{i+1/2,j}^{x,\Delta x}(\tau) - F_{i-1/2,j}^{x,\Delta x}(\tau)\right)\,\mathrm{d}\tau - \frac{1}{\Delta y}\int_t^{t+\Delta t}\left(F_{i,j+1/2}^{y,\Delta y}(\tau) - F_{i,j-1/2}^{y,\Delta y}(\tau)\right)\,\mathrm{d}\tau. \tag{18b}$$

We denote the evolution operator corresponding to (18) and associated with the two-dimensional mesh size $\Delta := (\Delta x, \Delta y)$ by $\mathcal{S}^{\Delta}$.

4.1.2 Weak Convergent Schemes*

The next ingredient in the construction of approximate EMV solutions for (11) is to employ the above numerical schemes in the following three-step algorithm.

Algorithm 5

Step 1: *Let $u_0 : \Omega \mapsto L^\infty(\mathbb{R}^d)$ be a random field on a probability space $(\Omega, \mathcal{F}, P)$ such that the initial Young measure σ in (11) is the law of the random field u_0 (see Proposition 2).*
Step 2: *We evolve the initial random field by applying the numerical scheme (17a) for every $\omega \in \Omega$ to obtain an approximation $u^{\Delta x}(\omega) := \mathcal{S}^{\Delta x}u_0(\omega)$ to the solution random field $u(\omega)$, corresponding to the initial random field $u_0(\omega)$.*
Step 3: *Define the approximate measure-valued solution $\nu^{\Delta x}$ as the law of $u^{\Delta x}$ with respect to P (in Appendix A.3.1).*

By Proposition 1 in Appendix A.3.1, $\nu^{\Delta x}$ is a Young measure. This sequence of Young measures $\nu^{\Delta x}$ serves as approximations to the EMV solutions of (11).

Next, we show that if the numerical scheme (17a) satisfies a set of criteria, then the approximate Young measures $\nu^{\Delta x}$ generated by Algorithm 4.1 will converge weak* to an EMV solution of (11). Specific examples for such weak* convergent schemes are provided in Sect. 5. To simplify the presentation, we restrict attention to the one-dimensional case; the argument is readily extended to the general multi-dimensional case, and the details can be found in [28] (see Sect. 3.2, in particular Lemmas 3.4 and 3.5).

Theorem 6 *Assume that the approximate solutions $u^{\Delta x}$ generated by the one-dimensional numerical scheme (17) satisfy the following.*

– *Uniform boundedness:*

$$\left\|u^{\Delta x}(\omega)\right\|_{L^\infty(\mathbb{R}\times\mathbb{R}_+)} \leqslant C, \quad \forall\,\omega\in\Omega,\ \Delta x > 0. \tag{19a}$$

– *Weak BV: There exists $1 \leqslant r < \infty$ such that*

$$\lim_{\Delta x\to 0}\int_0^T \sum_i \left|u_{i+1}^{\Delta x}(\omega,t) - u_i^{\Delta x}(\omega,t)\right|^r \Delta x\,dt = 0 \quad \forall\,\omega\in\Omega \tag{19b}$$

– *Entropy consistency: The numerical scheme (17a) is entropy stable with respect to an entropy pair (η, q) i.e., there exists a numerical entropy flux $Q = Q_{i+1/2}(t)$, consistent with the entropy flux q and locally Lipschitz, such that computed solutions satisfy the discrete entropy inequality*

$$\frac{d}{dt}\eta(u^{\Delta x}) + \frac{1}{\Delta x}\left(Q^{\Delta x}_{i+1/2} - Q^{\Delta x}_{i-1/2}\right) \leqslant 0 \quad \forall\, t > 0,\ i \in \mathbb{Z},\ \omega \in \Omega. \tag{19c}$$

– *Consistency with initial data: If $\sigma^{\Delta x}$ is the law of $u_0^{\Delta x}$, then*

$$\lim_{\Delta x \to 0} \int_{\mathbb{R}} \psi(x) \left\langle \sigma_x^{\Delta x}, \mathrm{id} \right\rangle \, dx = \int_{\mathbb{R}} \psi(x) \left\langle \sigma_x, \mathrm{id} \right\rangle \, dx \quad \forall\, \psi \in C_c^1(\mathbb{R}). \tag{19d}$$

and

$$\limsup_{\Delta x \to 0} \int_{\mathbb{R}} \psi(x) \left\langle \sigma_x^{\Delta x}, \eta \right\rangle \, dx \leqslant \int_{\mathbb{R}} \psi(x) \left\langle \sigma_x, \eta \right\rangle \, dx \quad \forall\, 0 \leqslant \psi \in C_c^1(\mathbb{R}) \tag{19e}$$

Then the approximate Young measures $\nu^{\Delta x}$ converge weak (up to a subsequence) as $\Delta x \to 0$, to an EMV solution $\nu \in \mathbf{Y}(\mathbb{R} \times \mathbb{R}_+, \mathbb{R}^N)$ of (11).*

Proof From the assumption (19a) that $u^{\Delta x}$ is L^∞-bounded, it follows that $\nu^{\Delta x}$ is uniformly bounded, in the sense that its support $\operatorname{supp} \nu^{\Delta x}_{(x,t)}$ lies in a fixed compact subset of $\mathbb{R}^N$ for every (x, t); see Appendix A.2.2. The fundamental theorem of Young measures (see Appendix A.2.6) gives the existence of a $\nu \in \mathbf{Y}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^N)$ and a subsequence of $\nu^{\Delta x}$ such that $\nu^{\Delta x} \rightharpoonup \nu$ weak*.

First, we show that the limit Young measure ν satisfies the entropy inequality (14). To this end, let $\varphi \in C_c^1(\mathbb{R} \times [0, T))$. Then

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^d} \left\langle \nu_{(x,t)}, \eta \right\rangle \partial_t \varphi(x,t) + \left\langle \nu_{(x,t)}, q \right\rangle \partial_x \varphi(x,t) \, \mathrm{d}x\mathrm{d}t \\
&\quad = \lim_{\Delta x \to 0} \int_0^T \int_{\mathbb{R}^d} \left\langle \nu^{\Delta x}_{(x,t)}, \eta \right\rangle \partial_t \varphi(x,t) + \left\langle \nu^{\Delta x}_{(x,t)}, q \right\rangle \partial_x \varphi(x,t) \, \mathrm{d}x\mathrm{d}t
\end{aligned}$$

by the weak* convergence $\nu^{\Delta x} \rightharpoonup \nu$. Denote $\eta^{\Delta x}(\omega, x, t) := \eta(u^{\Delta x}(\omega, x, t))$. Then for every $\Delta x > 0$ we have

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^d} \left\langle \nu^{\Delta x}_{(x,t)}, \eta \right\rangle \partial_t \varphi(x,t) \, \mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d} \varphi(x,0) \left\langle \sigma_x^{\Delta x}, \eta \right\rangle \mathrm{d}x \\
&= \int_{\mathbb{R}} \int_0^T -\partial_t \left\langle \nu^{\Delta x}_{(x,t)}, \eta \right\rangle \varphi(x,t) \, \mathrm{d}t\mathrm{d}x \\
&= \int_\Omega \int_{\mathbb{R}} \int_0^T -\partial_t \eta^{\Delta x}(\omega, x, t) \varphi(x,t) \, \mathrm{d}t\mathrm{d}x\mathrm{d}P(\omega) \\
&\geqslant \int_\Omega \int_{\mathbb{R}} \int_0^T \sum_i \mathbb{1}_{\mathcal{C}_i}(x) \frac{Q_{i+1/2}(\omega,t) - Q_{i-1/2}(\omega,t)}{\Delta x} \varphi(x,t) \mathrm{d}t\mathrm{d}x\mathrm{d}P(\omega)
\end{aligned}$$

$$
\begin{aligned}
&= \int_\Omega \int_0^T \sum_i \frac{Q_{i+1/2}(\omega,t) - Q_{i-1/2}(\omega,t)}{\Delta x} \int_{\mathcal{C}_i} \varphi(x,t)\,\mathrm{d}x\mathrm{d}t\mathrm{d}P(\omega) \\
&= \int_\Omega \int_0^T \sum_i \left(Q_{i+1/2}(\omega,t) - Q_{i-1/2}(\omega,t)\right) \overline{\varphi}_i^{\Delta x}(t)\,\mathrm{d}t\mathrm{d}P(\omega) \\
&= -\int_\Omega \int_0^T \sum_i Q_{i+1/2}(\omega,t) \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x}\,\Delta x\mathrm{d}t\mathrm{d}P(\omega) \\
&= -\int_\Omega \int_0^T \sum_i q(u_i^{\Delta x}(\omega,t)) \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x}\,\Delta x\mathrm{d}t\mathrm{d}P(\omega) \\
&\quad - \int_\Omega \int_0^T \sum_i \left(Q_{i+1/2}(\omega,t) - q(u_i^{\Delta x}(\omega,t))\right) \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x}\,\Delta x\mathrm{d}t\mathrm{d}P(\omega).
\end{aligned}
$$

(We have written $\overline{\varphi}_i^{\Delta x}(t) := \frac{1}{\Delta x}\int_{\mathcal{C}_i} \varphi(x,t)\,\mathrm{d}x$.) The first term can be written as

$$
\begin{aligned}
&-\int_\Omega \int_0^T \sum_i q\left(u_i^{\Delta x}(\omega,t)\right) \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x}\,\Delta x\mathrm{d}t \\
&\quad = -\int_0^T \sum_i \left\langle \nu_{(x_i,t)}^{\Delta x}, q\right\rangle \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x}\,\Delta x\mathrm{d}t\mathrm{d}P(\omega) \\
&\quad \to -\int_0^T \int_{\mathbb{R}} \left\langle \nu_{(x,t)}, q\right\rangle \partial_x \varphi(x,t)\,\mathrm{d}x\mathrm{d}t.
\end{aligned}
$$

The second term goes to zero:

$$
\begin{aligned}
&\left| \int_\Omega \int_0^T \sum_i \left(Q_{i+1/2}(\omega,t) - q(u_i^{\Delta x}(\omega,t))\right) \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x}\,\Delta x\mathrm{d}t\mathrm{d}P(\omega) \right| \\
&\quad \leqslant C \int_\Omega \int_0^T \sum_i \left|u_{i+1}^{\Delta x}(\omega,t) - u_i^{\Delta x}(\omega,t)\right| \left| \frac{\overline{\varphi}_{i+1}^{\Delta x}(t) - \overline{\varphi}_i^{\Delta x}(t)}{\Delta x} \right| \Delta x\mathrm{d}t\mathrm{d}P(\omega) \\
&\quad \leqslant C \sup_\omega \left(\int_0^T \sum_i \left|u_{i+1}^{\Delta x}(\omega,t) - u_i^{\Delta x}(\omega,t)\right|^r \Delta x\mathrm{d}t \right)^{1/r} \|\partial_x \varphi\|_{L^{r'}(\mathbb{R}\times(0,T))} \\
&\quad \to 0
\end{aligned}
$$

by (19b), where r' is the conjugate exponent of r. In conclusion, the limit ν satisfies (14).

The proof that the limit measure ν satisfies (12) follows from the above by setting $\eta = \pm\,\mathrm{id}$ and $q = \pm f$. □

A similar construction can be readily performed in several space dimensions. To this end, we replace $\mathcal{S}^{\Delta x}$ in Step 2 of Algorithm 5 with the two-dimensional solution

operator $\mathcal{S}^{\Delta}$, and the corresponding approximate solution $u^{\Delta x}$ with u^{Δ}. The weak* convergence of the resulting approximate Young measure ν^{Δ} is described below.

Theorem 7 *Assume that the approximate solutions u^{Δ} generated by scheme (18a) satisfy the following.*

– *Uniform boundedness:*

$$\|u^{\Delta}(\omega)\|_{L^{\infty}(\mathbb{R}^2\times\mathbb{R}_+)} \leqslant C, \quad \forall \omega \in \Omega, \Delta x, \Delta y > 0. \tag{20}$$

– *Weak BV: There exists $1 \leqslant r < \infty$ such that $\forall\ \omega \in \Omega$*

$$\lim_{\Delta x,\Delta y\to 0} \int_0^T \sum_{i,j} \left(\left| u^{\Delta}_{i+1,j}(\omega,t) - u^{\Delta}_{i,j}(\omega,t) \right|^r + \left| u^{\Delta}_{i,j+1}(\omega,t) - u^{\Delta}_{i,j}(\omega,t) \right|^r \right) \Delta x \Delta y dt = 0. \tag{21}$$

– *Entropy consistency: The numerical scheme (18a) is entropy stable with respect to an entropy pair (η, q), in the sense that there exist locally Lipschitz numerical entropy fluxes $(Q^{x,\Delta x}, Q^{y,\Delta y}) = (Q^{x,\Delta x}_{i+1/2,j}(t), Q^{y,\Delta y}_{i,j+1/2}(t))$, consistent with the entropy flux $q = (q^x, q^y)$, such that computed solutions satisfy the discrete entropy inequality, namely $\forall\ t > 0,\ i, j \in \mathbb{Z},\ \omega \in \Omega$*

$$\frac{d}{dt}\eta(u^{\Delta}) + \frac{1}{\Delta x}\left(Q^{x,\Delta x}_{i+1/2,j} - Q^{x,\Delta x}_{i-1/2,j}\right) + \frac{1}{\Delta y}\left(Q^{y,\Delta y}_{i,j+1/2} - Q^{y,\Delta y}_{i,j-1/2}\right) \leqslant 0. \tag{22}$$

– *Consistency with initial data: Let σ^{Δ} be the law of the random field u^{Δ}_0 that approximates the initial random field u_0. Then, the consistency conditions (19d) and (19e) hold.*

Then, the approximate Young measures ν^{Δ} converge weak (up to a subsequence) to a Young measure $\nu \in \mathbf{Y}(\mathbb{R}^2 \times \mathbb{R}_+, \mathbb{R}^N)$ as $\Delta x, \Delta y \to 0$ and ν is an EMV solution of (11), i.e.,*

The proof of the above theorem is a simple generalization of the proof of convergence theorem 6, see Section 3.2 of [28] (in particular Lemmas 3.4 and 3.5) for details. The above construction can also be readily extended to three spatial dimensions.

Remark 4 The uniform L^{∞} bound (19a), (20) is a technical assumption that we require in this article. This assumption can be relaxed to only an L^p bound. This extension is described in a forthcoming paper [29].

Remark 5 The conditions (19d) and (19e), which say that $\sigma^{\Delta x} \to \sigma$ in a certain sense, are weaker than weak* convergence. It is readily checked that a sufficient condition for this is that $u_0 \in L^1(\mathbb{R}; \mathbb{R}^N) \cap L^{\infty}(\mathbb{R}; \mathbb{R}^N)$ and $u_0^{\Delta x}(\omega, \cdot) \to u_0(\omega, \cdot)$ in $L^1(\mathbb{R}^d; \mathbb{R}^N)$ for all $\omega \in \Omega$ (which in fact implies that $\sigma^{\Delta x} \to \sigma$ strongly).

4.1.3 Weak Convergence with Atomic Initial Data*

In view of the non-uniqueness example 1, one can not expect an unique construction of EMV solutions for general MV initial data. Instead, as argued before, we focus our attention on perturbations of atomic initial data $\sigma = \delta_{u_0}$ for some $u_0 \in L^1(\mathbb{R}^d, \mathbb{R}^N) \cap L^\infty(\mathbb{R}^d, \mathbb{R}^N)$. We construct approximate EMV solutions of (11) in this case using the following specialization of Algorithm 4.1.

Algorithm 8 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X : \Omega \to L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ be a random variable satisfying $\|X\|_{L^1(\mathbb{R}^d)} \leqslant 1$ P-almost surely.*

Step 1 *Fix a small number $\varepsilon > 0$. Perturb u_0 by defining $u_0^\varepsilon(\omega, x) := u_0(x) + \varepsilon X(\omega, x)$. Let σ^ε be the law of u_0^ε.*
Step 2 *For each $\omega \in \Omega$, let $u^{\Delta x,\varepsilon}(\omega) := \mathcal{S}^{\Delta x} u_0^\varepsilon(\omega)$, with $\mathcal{S}^{\Delta x}$ being the solution operator corresponding to the numerical scheme* (17).
Step 3 *Let $\nu^{\Delta x,\varepsilon}$ be the law of $u^{\Delta x,\varepsilon}$ with respect to P.* □

Theorem 9 *Let $\{\nu^{\Delta x,\varepsilon}\}$ be the family approximate EMV solutions constructed by Algorithm 8. Then there exists a subsequence $(\Delta x_n, \varepsilon_n) \to 0$ such that*

$$\nu^{\Delta x_n,\varepsilon_n} \rightharpoonup \nu \in \mathcal{E}(\delta_{u_0}),$$

that is, $\nu^{\Delta x_n,\varepsilon_n}$ converges weak to an EMV solution ν with atomic initial data u_0.*

Proof By Theorem 6, we know that for every $\varepsilon > 0$ there exists a subsequence $\nu^{\Delta x_n,\varepsilon}$ which converges weak* to an EMV solution ν^ε of (11) with initial data σ^ε. Thus, (14) holds with (ν, σ) replaced by $(\nu^\varepsilon, \sigma^\varepsilon)$; we abbreviate the corresponding entropy statement as $(14)_\varepsilon$. The convergence of the sequence ν^{ε_n} as $\varepsilon_n \to 0$ is a consequence of the fundamental theorem of Young measures: By Theorem 13, there exists a weak* convergent subsequence $\nu^{\varepsilon_n} \rightharpoonup \nu$. The fact that ν is an EMV solution follows at once by taking the limit $\varepsilon_n \to 0$ in $(14)_{\varepsilon_n}$. □

4.2 What are We Computing? Weak* Convergence of Space–Time Averages

We begin by quoting [48, p. 143]: "Just because we cannot prove that compressible flows with prescribed initial values exist doesn't mean that we cannot compute them." The question is what are the computed quantities encoded in the EMV solutions.

According to Theorems 6, 9, the approximations generated by Algorithm 4.1 and 4.5 converge to an EMV solution in the following sense: For all $g \in C_0(\mathbb{R}^N)$ and $\psi \in L^1(\mathbb{R}^d \times \mathbb{R}_+)$,

$$\lim_{\Delta x \to 0} \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \psi(x,t) \left\langle \nu^{\Delta x}_{(x,t)}, g \right\rangle \mathrm{d}x\mathrm{d}t = \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \psi(x,t) \left\langle \nu_{(x,t)}, g \right\rangle \mathrm{d}x\mathrm{d}t. \quad (23)$$

As we assume that the approximate solutions are L^∞-bounded (property (19a)), any $g \in C(\mathbb{R}^N)$ can serve as a test function in (23), see Appendix A.2.6. In particular, we can choose $g(\xi) = \xi$ to obtain the *mean* of the measure-valued solution. Similarly,

the variance can be computed by choosing the test function $g(\xi) = \xi \otimes \xi$. Higher statistical moments can be computed analogously.

In practice, the goal of any numerical simulation is to accurately compute *statistics of space–time averages* or *statistics of functionals of interest* of solution variables and to compare them to experimental or observational data. Thus, the weak* convergence of approximate Young measures, computed by Algorithms 4.1 and 4.5, provides an approximation of exactly these *observable* quantities of interest.

4.2.1 Monte Carlo Approximation

In order to compute statistics of space–time averages in (23), we need to compute phase space integrals with respect to the measure $\nu^{\Delta x}$:

$$\left\langle \nu^{\Delta x}_{(x,t)}, g \right\rangle := \int_{\mathbb{R}^N} g(\xi)\, \mathrm{d}\nu^{\Delta x}_{(x,t)}(\xi).$$

The last ingredient in our construction of EMV solutions, therefore, is numerical approximation which is necessary to compute these phase space integrals. To this end, we utilize the equivalent representation of the measure $\nu^{\Delta x}$ as the law of the *random field* $u^{\Delta x}$:

$$\left\langle \nu^{\Delta x}_{(x,t)}, g \right\rangle := \int_{\mathbb{R}^N} g(\xi)\, \mathrm{d}\nu^{\Delta x}_{(x,t)}(\xi) = \int_{\Omega} g\left(u^{\Delta x}(\omega; x, t)\right)\, \mathrm{d}P(\omega). \tag{24}$$

We will approximate this integral by a Monte Carlo sampling procedure:

Algorithm 10 *Let $\Delta x > 0$ and let M be a positive integer. Let $\sigma^{\Delta x}$ be the initial Young measure in (11) and let $u_0^{\Delta x}$ be a random field $u_0^{\Delta x} : \Omega \times \mathbb{R}^d \to \mathbb{R}^N$ such that $\sigma^{\Delta x}$ is the law of $u_0^{\Delta x}$.*

Step 1 *Draw M independent and identically distributed random fields $u_0^{\Delta x,k}$ for $k = 1, \ldots, M$.*
Step 2 *For each k and* for a fixed $\omega \in \Omega$, *use the finite difference scheme (17a) to numerically approximate the conservation law (1) with initial data $u_0^{\Delta x,k}(\omega)$. Denote $u^{\Delta x,k}(\omega) = \mathcal{S}^{\Delta x} u_0^{\Delta x,k}(\omega)$.*
Step 3 *Define the approximate measure-valued solution $\nu^{\Delta x,M} := \frac{1}{M}\sum_{k=1}^{M} \delta_{u^{\Delta x,k}(\omega)}$.*

For every $g \in C(\mathbb{R}^N)$, we have

$$\left\langle \nu^{\Delta x,M}, g \right\rangle = \frac{1}{M}\sum_{k=1}^{M} g\left(u^{\Delta x,k}(\omega)\right).$$

Thus, the space–time average (23) is approximated by

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)\left\langle \nu^{\Delta x}_{(x,t)},g\right\rangle \,\mathrm{d}x\mathrm{d}t \approx \frac{1}{M}\sum_{k=1}^{M}\int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)g\big(u^{\Delta x,k}(\omega;x,t)\big)\,\mathrm{d}x\mathrm{d}t. \tag{25}$$

Note that, as in any Monte Carlo method, the approximation $\nu^{\Delta x,M}$ depends on the choice of $\omega\in\Omega$, i.e., the choice of seed in the random number generator. However, we can prove that the quality of approximation is independent of this choice, P-almost surely:

Theorem 11 (Convergence for large samples) *Algorithm 10 converges, that is,*

$$\nu^{\Delta x,M}\rightharpoonup\nu^{\Delta x}\quad \textit{weak*},$$

and, for a subsequence $M\to\infty$, P-almost surely. Equivalently, for every $\psi\in L^1(\mathbb{R}^d\times\mathbb{R}_+)$ and $g\in C(\mathbb{R}^N)$,

$$\lim_{M\to\infty}\frac{1}{M}\sum_{k=1}^{M}\int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)g\big(u^{\Delta x,k}(x,t)\big)dxdt = \int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)\left\langle\nu^{\Delta x}_{(x,t)},g\right\rangle dxdt. \tag{26}$$

The limits are uniform in Δx.

The proof involves an adaptation of the law of large numbers for the present setup and is provided in Appendix 2. Combining (26) with the convergence established in Theorem 6, we conclude with the following.

Corollary 1 (Convergence with mesh refinement) *There are subsequences $\Delta x\to 0$ and $M\to\infty$ such that*

$$\nu^{\Delta x,M}\rightharpoonup\nu\quad \textit{weak*},$$

or equivalently, for every $\psi\in L^1(\mathbb{R}^d\times\mathbb{R}_+)$ and $g\in C(\mathbb{R}^N)$,

$$\lim_{\Delta x\to 0}\lim_{M\to\infty}\frac{1}{M}\sum_{k=1}^{M}\int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)g\big(u^{\Delta x,k}(x,t)\big)\,dxdt = \int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)\left\langle\nu_{(x,t)},g\right\rangle dxdt \tag{27}$$

The limits in Δx and M are interchangeable.

5 Examples of Weak* Convergent Numerical Schemes

In this section, we provide concrete examples of numerical schemes that satisfy the criteria (19) of Theorem 6, for weak* convergence to EMV solutions of (11). We start with scalar conservation laws.

5.1 Scalar Conservation Laws

We begin by considering scalar conservation laws. Monotone finite difference (volume) schemes (see [15,37] for a precise definition) for scalar equations are uniformly bounded in L^∞ (as they satisfy a discrete maximum principle), satisfy a discrete entropy inequality (using the Crandall-Majda numerical entropy fluxes [15]) and are TVD—the total variation of the approximate solutions is non-increasing over time. Consequently, the approximate solutions satisfy the weak BV estimate (19b) (resp, (21) in the multi-dimensional case) with $r = 1$. Thus, monotone schemes, approximating scalar conservation laws, satisfy all the abstract criteria of Theorem 6.

In fact, one can obtain a precise convergence rate for monotone schemes [46]:

$$\left\| u^{\Delta x}(\omega,\cdot,t) - u(\omega,\cdot,t) \right\|_{L^1(\mathbb{R}^d)} \leqslant Ct\,\mathrm{TV}(u_0(\omega))\sqrt{|\Delta x|} \quad \forall\,\omega, \tag{28}$$

where $u(\omega) = \lim_{\Delta x\to 0} u^{\Delta x}(\omega)$ denotes the entropy solution of the Cauchy problem for a scalar conservation law with initial data $u_0(\omega)$. Using this error estimate, we obtain the following strong convergence results for monotone schemes.

Theorem 12 *Let $\nu^{\Delta x}$ be generated by Algorithm 5, and let ν be the law of the entropy solution $u(\omega)$. If $\mathrm{TV}(u_0(\omega)) \leqslant C$ for all $\omega \in \Omega$, then $\nu^{\Delta x} \to \nu$ strongly as $\Delta x \to 0$.*

Proof Define $\pi_z^{\Delta x} \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^N)$ as the law of the random variable $\left(u^{\Delta x}(z), u(z)\right)$,

$$\pi_z^{\Delta x}(A) := P\left(\left(u^{\Delta x}(z), u(z)\right) \in A\right), \quad A \subset \mathbb{R}\times\mathbb{R} \text{ Borel measurable.}$$

Then $\pi_z^{\Delta x}$ is a Young measure for all z and $\Delta x > 0$. Clearly, $\pi_z^{\Delta x} \in \Pi\left(\nu_z^{\Delta x}, \nu_z\right)$, and hence

$$W_1\left(\nu_z^{\Delta x}, \nu_z\right) \leqslant \int_{\mathbb{R}^N\times\mathbb{R}^N} |\xi-\zeta|\, d\pi_z^{\Delta x}(\xi,\zeta) = \int_\Omega |u^{\Delta x}(\omega,x,t) - u(\omega,x,t)|\, \mathrm{d}P(\omega).$$

Hence, by Kutznetsov's error estimate (28),

$$\int_0^T \int_{\mathbb{R}} W_1\left(\nu_z^{\Delta x}, \nu_z\right) \mathrm{d}x\mathrm{d}t \leqslant C\sqrt{|\Delta x|} \to 0 \quad \text{as } \Delta x \to 0.$$

□

Remark 6 We can relax the uniform boundedness of $\mathrm{TV}(u_0(\omega))$ to just integrability of the function $\omega \mapsto \mathrm{TV}(u_0(\omega))$.

Remark 7 Note that, in light of Theorem 1 and Example 1, the limit entropy measure-valued solution ν is unique only if the initial measure-valued data σ is atomic.

5.2 Systems of Conservation Laws

The convergence theorems 6 and 7 specify that the ensemble-based algorithms converge if the underlying numerical scheme satisfies the criteria (19). We recall from remark 4 that the L^∞ bound (19a), (20), is a technical assumption that cannot be rigorously verified for any of the existing numerical schemes approximating systems of conservation laws. Although this assumption holds for many numerical examples (for instance in all the numerical experiments presented in the next section), it is unclear whether this bound holds for systems of conservation laws with general L^∞ initial data. However, this assumption can be relaxed to requiring uniform L^p bounds on numerical schemes with $1 \leqslant p < \infty$, if the underlying systems of conservation laws possess a strictly convex entropy. Then, the solution framework utilizes generalized Young measures that allow for concentrations as well as persistent oscillations in sequences of approximations. Consequently, any entropy stable scheme for conservation laws with strictly convex entropy will satisfy an uniform L^2 bound and will converge to the corresponding (generalized) entropy measure-valued solution. Such an extension is presented in a forthcoming paper [29].

Hence, we will present numerical schemes for approximating systems of conservation laws that satisfy the weak BV bounds (19b), (21) and the discrete entropy inequalities (19c). We present examples of such schemes below.

5.2.1 The ELW Scheme

The **ELW** (entropy stable Lax–Wendroff) scheme, introduced in [28, Section 4.2], is finite difference scheme of the form (17a) with flux function

$$F_{i+1/2} := \tilde{F}^p_{i+1/2} - c_{i+1/2} \left| [\![v]\!]_{i+1/2} \right|^{p-1} [\![v]\!]_{i+1/2} \tag{29}$$

where $\tilde{F}^p_{i+1/2}$ is a pth order accurate ($p \in \mathbb{N}$) entropy conservative numerical flux (see [49,61]), $c_{i+1/2}$ is some positive number, $[\![v]\!]_{i+1/2} := v_{i+1} - v_i$ and $v := \eta'(u)$ is the entropy variable.

This scheme was shown to be (formally) p-th order accurate, entropy stable, satisfies the weak BV bound (19b), and converges pointwise for scalar conservation laws [28, Proposition 4.2 and 4.3]. The two-dimensional version of the scheme, with the corresponding discrete entropy inequality and weak BV bound (21), is straightforward to construct. Hence, under the assumptions (19a),(20) that the scheme is bounded in L^∞, the approximate measure-valued solutions generated by the ELW scheme converge to an entropy measure-valued solution of (11).

5.2.2 TeCNO Finite Difference Schemes

The **TeCNO** schemes, introduced in [27,28], are finite difference schemes of the form (17a) with flux function

$$F_{i+1/2} := \tilde{F}^p_{i+1/2} - \frac{1}{2} D_{i+1/2}\big(v^-_{i+1} - v^+_i\big). \tag{30}$$

Here, $\tilde{F}^p_{i+1/2}$ is an entropy conservative flux as in (29), $D_{i+1/2}$ is a positive definite matrix, and $v^\pm_j$ are the cell interface values of a p-th order accurate ENO reconstruction of the entropy variable v (see [26,39]). The multi-dimensional (Cartesian) version was also designed in [27], see also [28]. It was shown in [27,28] that the TeCNO schemes

- are (formally) p-th order accurate
- are entropy stable—they satisfy a discrete entropy inequality of the form (19c) (see Theorem 4.1 of [27] for the one-dimensional case and Theorem 6.1 of [27] for the multi-dimensional case)
- have weakly bounded variation when $p = 1$ and $p = 2$, i.e., they satisfy a bound of the form (19b) in the one-dimensional case and (21) in two dimensions (see Theorem 6.6 of [28] and in general section 3.2 of [28] for the multi-dimensional case).

The weak total variation bound for $p \geqslant 3$ depends on a conjectured result for the ENO reconstruction method which remains to be proven (see [28, Section 5.5]). In light of these properties, and under the assumption (19a) that the scheme is bounded in L^∞, the approximate measure-valued solutions generated by the TeCNO scheme converge to an EMV solution of (11).

5.2.3 Shock Capturing Space–Time Discontinuous Galerkin (DG) Schemes

Although suitable for Cartesian grids, finite difference schemes of the type (17a) are difficult to extend to unstructured grids in several space dimensions. For problems with complex domain geometry that necessitates the use of unstructured grids (triangles, tetrahedra), an alternative discretization procedure is the space–time discontinuous finite element procedure of [4,40,41,43]. In this procedure, the entropy variables serve as degrees of freedom and entropy stable numerical fluxes like (30) need to be used at cell interfaces. Further stabilization terms like streamline diffusion and shock capturing terms are also necessary. In a recent paper [40], it was shown that a shock capturing streamline diffusion space–time DG method satisfied a discrete entropy inequality and a suitable version of the weak BV bound (19b), see Theorem 3.1 of [40] for the precise statements and results. Hence, this method was also shown to converge to an EMV solution in [40] (see Theorems 4.1 and 4.2 of [40]). We remark that the space–time DG methods are fully discrete, in contrast to semi-discrete finite difference schemes such as (17a).

6 Numerical Results

Our overall goal in this section will be to compute approximate EMV solutions of (11) with atomic initial data using Algorithm 8, as well as to investigate the stability of these solutions with respect to initial data. In Sects. 6.1 and 6.2, we consider the Kelvin–Helmholtz problem (8). In Section 6.3, we consider the Richtmeyer–Meshkov problem, see, e.g., [36] and the references therein.

For the rest of the section, we will present numerical experiments for the two-dimensional compressible Euler equations

$$\frac{\partial}{\partial t}\begin{pmatrix}\rho\\ \rho w^1\\ \rho w^2\\ E\end{pmatrix}+\frac{\partial}{\partial x_1}\begin{pmatrix}\rho w^1\\ \rho(w^1)^2+p\\ \rho w^1 w^2\\ (E+p)w^1\end{pmatrix}+\frac{\partial}{\partial x_2}\begin{pmatrix}\rho w^2\\ \rho w^1 w^2\\ \rho(w^2)^2+p\\ (E+p)w^2\end{pmatrix}=0. \tag{31}$$

Here, the density ρ, velocity field (w^1, w^2), pressure p and total energy E are related by the equation of state

$$E=\frac{p}{\gamma-1}+\frac{\rho\left((w^1)^2+(w^2)^2\right)}{2}.$$

The relevant entropy pair is given by

$$\eta(u)=\frac{-\rho s}{\gamma-1},\quad q^1(u)=w^1\eta(u),\quad q^2(u)=w^2\eta(u).$$

with $s=\log(p)-\gamma\log(\rho)$ being the thermodynamic entropy. The adiabatic constant γ is set to 1.4.

6.1 Kelvin–Helmholtz Problem: Mesh Refinement ($\Delta x\downarrow 0$)

As our first numerical experiment, we consider the two-dimensional compressible Euler equations of gas dynamics (31) with the initial data

$$u_0(x,\omega)=\begin{cases}u_L & \text{if } I_1<x_2<I_2\\ u_R & \text{if } x_2\leqslant I_1 \text{ or } x_2\geqslant I_2,\end{cases}\qquad x\in[0,1]^2 \tag{32}$$

with $\rho_L=2$, $\rho_R=1$, $w_L^1=-0.5$, $w_R^1=0.5$, $w_L^2=w_R^2=0$ and $p_L=p_R=2.5$.

The computational domain is $[0,1]^2$, and we consider periodic boundary conditions. Furthermore, the interface profiles

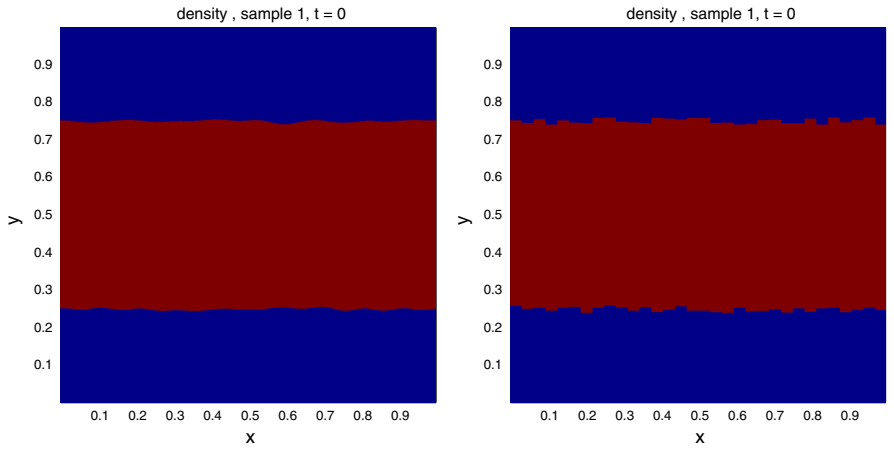


Fig. 5 Representative initial data for the Kelvin–Helmholtz problem: sinusoidal (*left*) and discontinuous (*right*) phase perturbation with $\varepsilon = 0.01$

$$I_j = I_j(x,\omega) := J_j + \varepsilon Y_j(x,\omega), \quad j = 1,2$$

are chosen to be small perturbations around $J_1 := 0.25$ and $J_2 := 0.75$, respectively, with

$$Y_j(x,\omega) = \sum_{n=1}^{m} a_j^n(\omega)\cos\left(b_j^n(\omega) + 2n\pi x_1\right), \quad j = 1,2.$$

Here, $a_j^n = a_j^n(\omega) \in [0,1]$ and $b_j^n = b_j^n(\omega) \in [-\pi,\pi]$, $i = 1,2$, $n = 1,\ldots,m$ are randomly chosen numbers. The coefficients a_j^n have been normalized such that $\sum_{n=1}^m a_j^n = 1$ to guarantee that $|I_j(x,\omega) - J_j| \leqslant \varepsilon$ for $j = 1,2$. We set $m = 10$.

Observe that by making ε small, this ω-ensemble of initial data lies inside an arbitrarily small ball centered at u_0. Indeed, it is readily checked that measured in, say, the $L^p([0,1]^2)$-norm, every sample $u_0(\cdot,\omega)$ is $O(\varepsilon^{1/p})$ away from the unperturbed steady state in (8).

A representative (single realization with fixed ω) initial datum for the density is shown in Fig. 5 (left). We observe that the resulting measure-valued Cauchy problem involves a random perturbation of the interfaces between the two streams (jets). This should be contrasted to initial value problem (8), (5), where the amplitude was randomly perturbed. We note that the law of the above initial datum can readily be written down and serves as the initial Young measure in the measure-valued Cauchy problem (11). Observe that this Young measure is non-atomic for some points in the domain.

6.1.1 Lack of Sample Convergence

We approximate the above MV Cauchy problem with the second-order entropy stable TeCNO2 scheme of [27]. In Fig. 6, we show the density at time $t = 2$ for a single

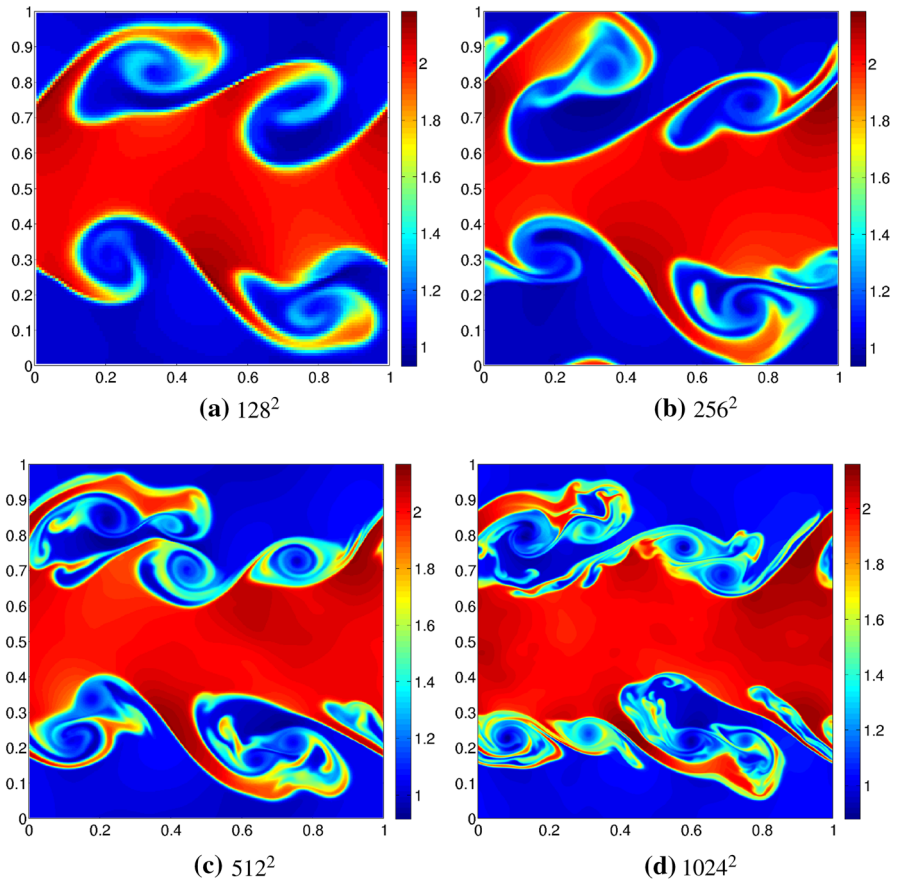


Fig. 6 Approximate density for the Euler equations (31) with initial data (32), $\varepsilon = 0.01$ and for a fixed ω (single sample), computed with the second-order TeCNO2 scheme of [27], at time $t = 2$ at different mesh resolutions

sample, i.e., for a fixed $\omega \in \Omega$, at different grid resolutions, ranging from 128^2 points to 1024^2 points. The figure suggests that the approximate solutions do not seem to converge as the mesh is refined. In particular, finer and finer-scale structures are formed as the mesh is refined, as already shown in Fig. 3. To further verify this lack of convergence, we compute the L^1 difference of the approximate solutions at successive mesh levels (7) and present the results in Fig. 7. We observe that this difference does not go to zero, suggesting that the approximate solutions may not converge as the mesh is refined.

6.1.2 Convergence of the Mean and Variance

The lack of convergence of the numerical schemes for single samples is not unexpected, given the results already mentioned in the introduction. Next, we will compute statistical quantities of the interest for this problem. First, we compute the Monte Carlo

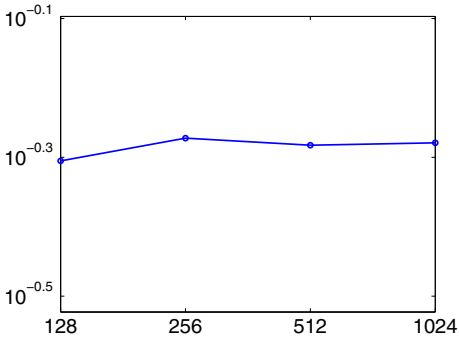


Fig. 7 Cauchy rates (7) at $t = 2$ for the density (y-axis) for a single sample of the Kelvin–Helmholtz problem, versus different mesh resolutions (x-axis)

approximation of the mean (25), denoted by $\bar{u}^{\Delta x}(x, t)$, at every point (x, t) in the computational domain. This sample mean of the density computed with $M = 400$ samples and the second-order TeCNO2 scheme is presented in Fig. 8 for a set of successively refined grid resolutions. The figure clearly suggests that the sample mean converges as the mesh is refined. This stands in stark contrast to the lack of convergence, at the level of single samples, as shown in Figs. 3 and 6. Furthermore, Fig. 8 also shows that small-scale structures, present in single sample (realization) computations, are indeed smeared or averaged out in the mean. This convergence of the mean is further quantified by computing the L^1 difference of the mean,

$$\|\bar{u}^{\Delta x} - \bar{u}^{\Delta x/2}\|_{L^1([0,1]^2)}. \tag{33}$$

and plotting the results in Fig. 9a. As predicted by the theory presented in Theorems 7 and 9, these results confirm that the sequence of approximate means forms a Cauchy sequence and hence converges to a limit as the mesh is refined. Similar convergence results were also observed for the means of the other conserved variables, namely momentum and total energy (not shown here). Furthermore, Fig. 8 also shows that the mean is varying in the y-direction only. This is completely consistent with the symmetries of the equations, of the initial data and the fact that periodic boundary conditions are employed. This is also in sharp contrast to the situation for single realizations where there is strong variation along both directions (see Fig. 6).

Next, we compute the sample variance and show the results in Fig. 10. The results suggest that the variance also converges with grid resolution. This convergence is also demonstrated quantitatively by plotting the L^1 differences of the variance at successive levels of resolution, shown in Fig. 9b. Again, the figure suggests that the sequence forms a Cauchy sequence and hence is convergent. Furthermore, the variance itself shows no small-scale features, even on very fine mesh resolutions (see Fig. 10). This figure also reveals that the variance is higher near the initial mixing layer.

6.1.3 Strong Convergence to an EMV Solution

Convergence of the mean and variance (as well as higher moments (not shown here)) confirm the weak* convergence predicted by (the multi-dimensional version of) The-

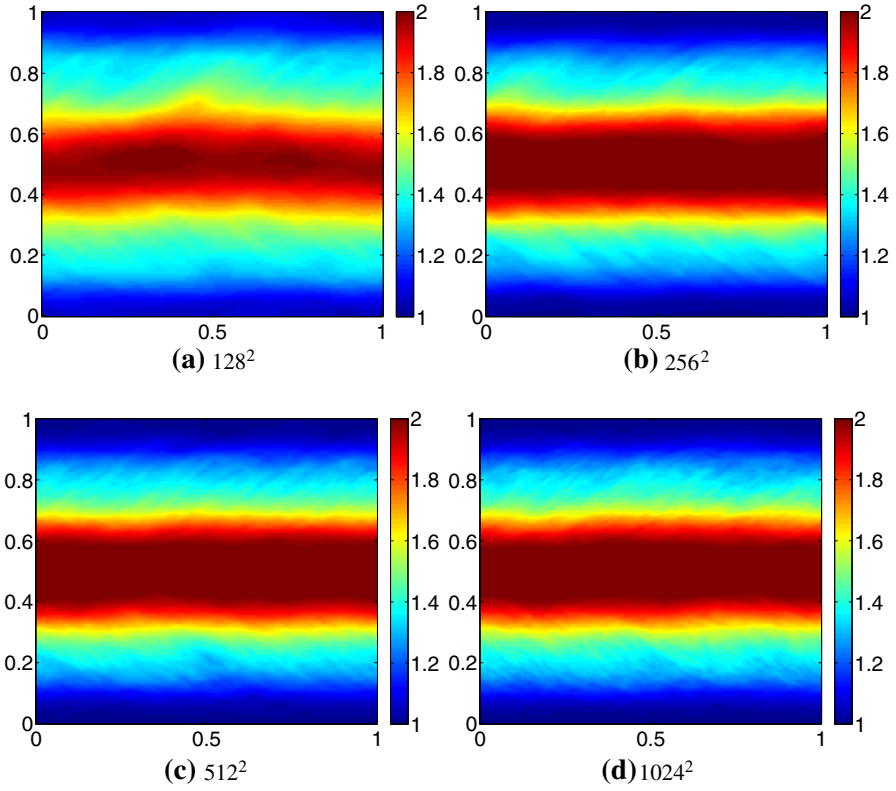


Fig. 8 Approximate sample means of the density for the Kelvin–Helmholtz problem (32) at time $t = 2$ and different mesh resolutions. All results are with 400 Monte Carlo samples

orems 6 and 9. Note that the convergence illustrated in Fig. 9 is in L^1 of the spatial domain. Next, we test strong convergence of the numerical approximations by computing the Wasserstein distance between two successive mesh resolutions:

$$W_1\left(\nu^{\Delta x}_{(x,t)}, \nu^{\Delta x/2}_{(x,t)}\right) \tag{34}$$

(see Appendix A.1.4). In Fig. 11, we show the L^1-norm of the Wasserstein distance between successive mesh resolutions

$$\left\| W_1\left(\nu^{\Delta x}_{(\cdot,t)}, \nu^{\Delta x/2}_{(\cdot,t)}\right) \right\|_{L^1([0,1]^2)} \tag{35}$$

at time $t = 2$; recall that this is the quantity appearing in (10). The figure suggests that this difference between successive mesh resolutions converges to zero. Hence, the approximate Young measures converge strongly in both space–time and phase space to a limit Young measure.

In Fig. 12, we show the pointwise difference in Wasserstein distance (35) between two successive mesh levels. The figure reveals that this distance decreases as the mesh

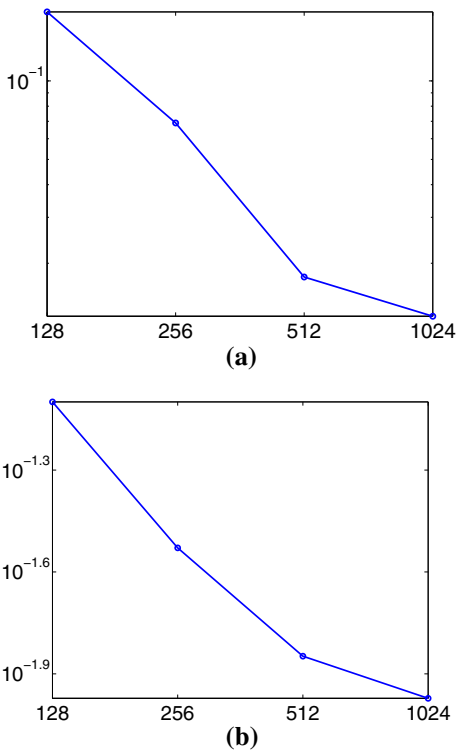


Fig. 9 Cauchy rates (33) for the sample mean and variance of the density (y-axis) versus mesh resolution (x-axis) for the Kelvin–Helmholtz problem (32). **a** Mean, **b** variance

is refined. Moreover, we see that the Wasserstein distance between approximate Young measures at successive resolutions is concentrated at the interface mixing layers. This is to be expected as the variance is also concentrated along these layers (cf. the variance plots in Fig. 10).

6.2 Kelvin–Helmholtz: Vanishing Variance Around Atomic Initial Data ($\varepsilon \downarrow 0$)

Our aim is to compute the entropy measure-valued solutions of the two-dimensional Euler equations with atomic initial measure, concentrated on the Kelvin–Helmholtz data (8). We utilize Algorithm 8 for this purpose and consider the perturbed initial data (32). Observe that this perturbed initial data converges strongly (cf. (10)) to the initial data (8) as $\varepsilon \to 0$. Following Algorithm 8, we wish to study the limit behavior of approximate solutions $\nu^{\Delta x,\varepsilon}$ as $\varepsilon \to 0$. To this end, we retain the same set up as the previous numerical experiment and compute approximate solutions using the TeCNO2 scheme of [27] at a very fine mesh resolution of 1024^2 points for different values of ε.

Results *for a single sample* at time $t = 2$ and different ε's are presented in Fig. 13. The figures indicate that there is no convergence as $\varepsilon \to 0$. The spread of the mixing region seems to remain large even when the perturbation parameter is reduced. This lack of convergence is further quantified in Fig. 14, where we plot the L^1 difference of

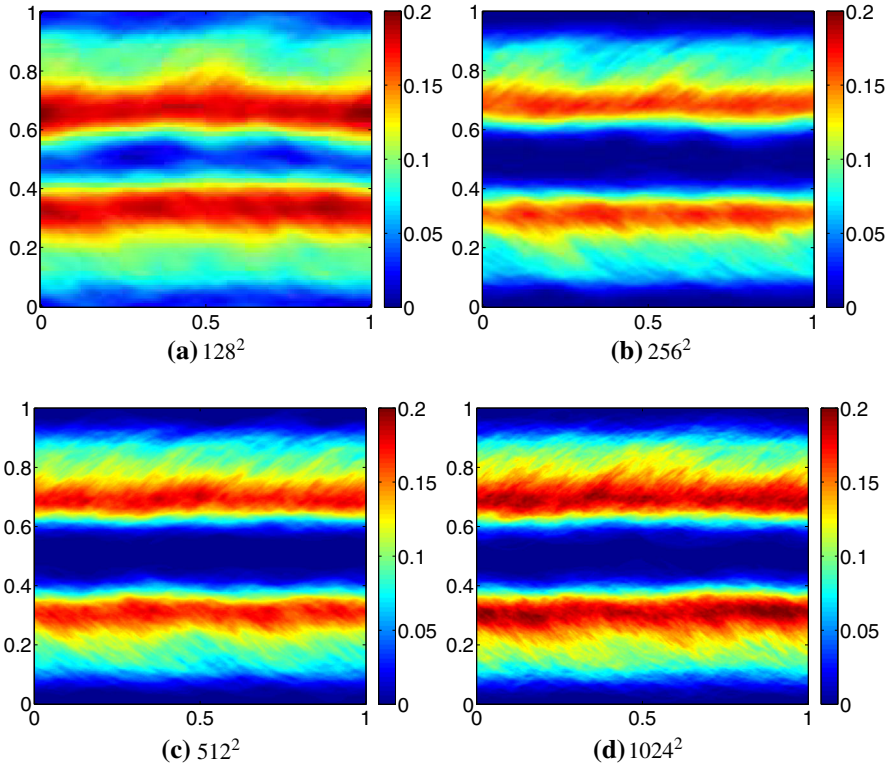


Fig. 10 Approximate sample variances of the density for the Kelvin–Helmholtz problem (32) at time $t = 2$ and different mesh resolutions. All results are with 400 Monte Carlo samples

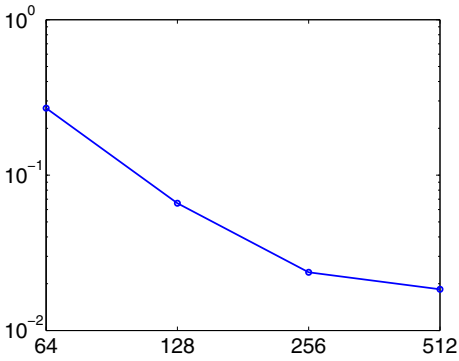
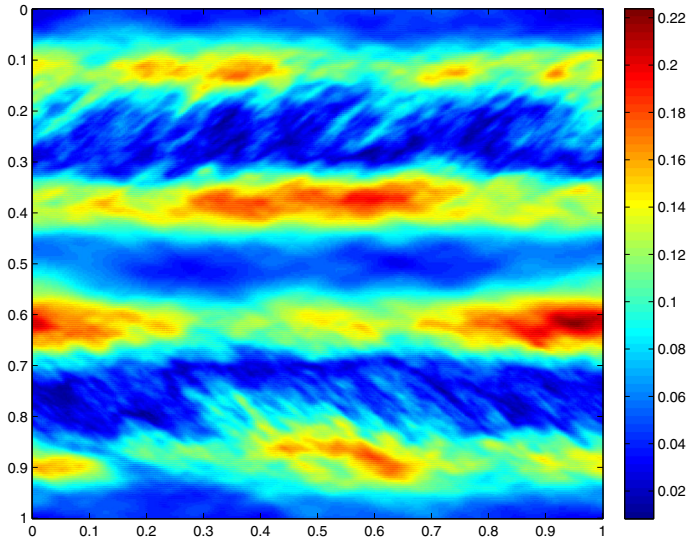


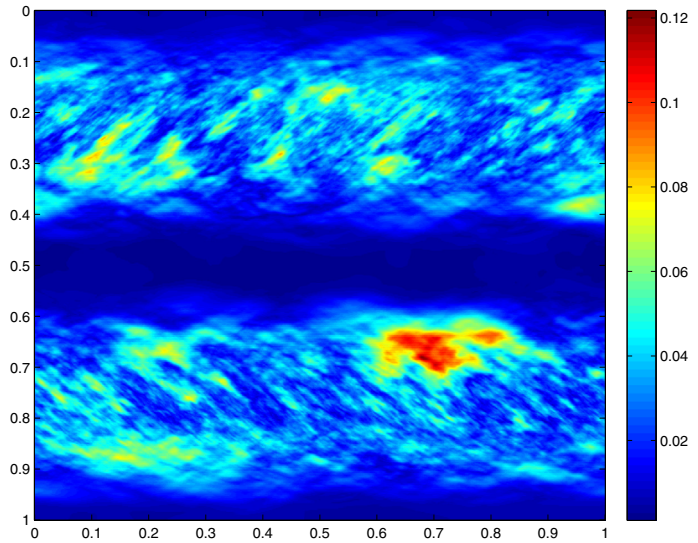
Fig. 11 Cauchy rates in the Wasserstein distance (35) at time $t = 2$ for the density (y-axis) with respect to different mesh resolutions (x-axis), for the Kelvin–Helmholtz problem (32)

the approximate density for successively reduced values of ε. This difference remains large even when ε is reduced by an order of magnitude.

Next, we compute the mean of the density over 400 samples at a fixed grid resolution of 1024^2 points and for different values of the perturbation parameter ε. This sample mean is plotted in Fig. 15. The figure clearly shows pointwise convergence as $\varepsilon \to 0$, to a limit different from the steady-state solution (8). This convergence of the mean with



(a) $W_1\left(\nu^{256}_{(x,t)}, \nu^{512}_{(x,t)}\right)$



(b) $W_1\left(\nu^{512}_{(x,t)}, \nu^{1024}_{(x,t)}\right)$

Fig. 12 Wasserstein distances between the approximate Young measure (density) (34) at successive mesh resolutions, at time $t = 2$

respect to decaying ε is quantified in Fig. 16a, where we compute the L^1 difference of the mean for successive values of ε. We observe that the mean forms a Cauchy sequence and hence converges.

Similarly the computations of the sample variance for different values of ε are presented in Fig. 17. Note that this figure, as well as the computations of the difference

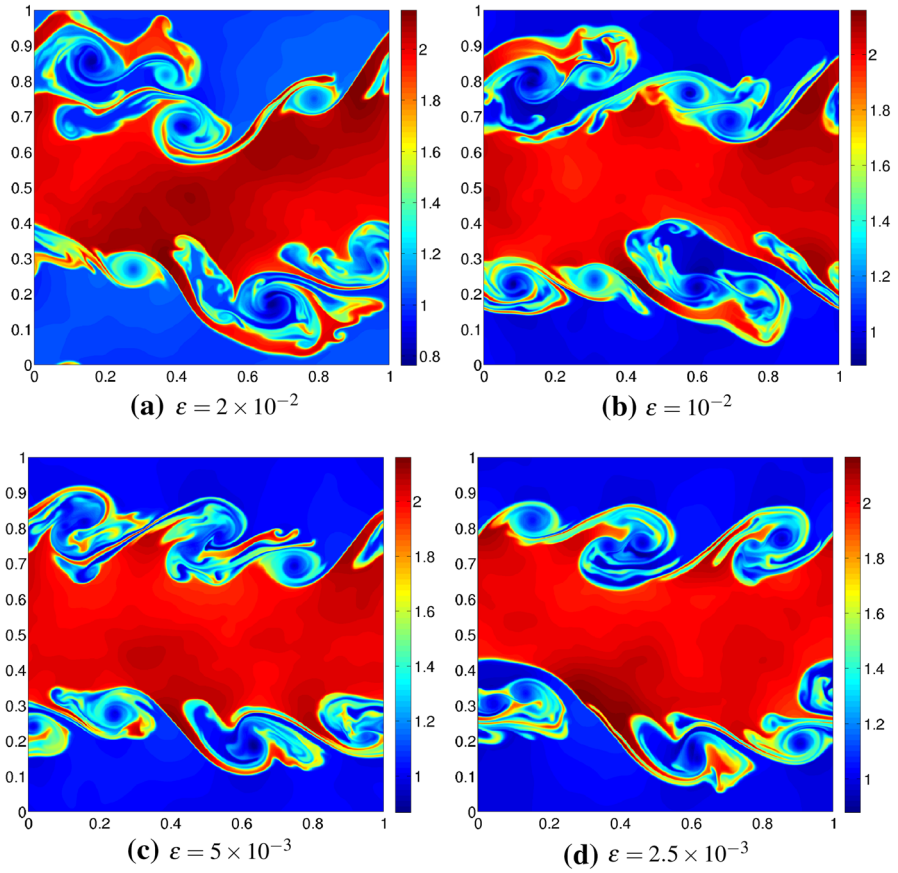


Fig. 13 Approximate density, computed with the TeCNO2 scheme for a single sample with initial data (32) for different initial perturbation amplitudes ε on a grid of 1024^2 points

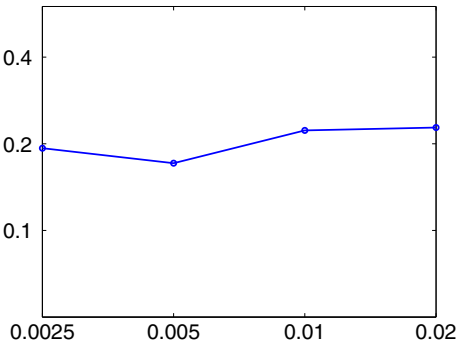


Fig. 14 Cauchy rates (L^1 difference for successively reduced ε) for the density (y-axis) at $t = 2$ for a single sample versus different values of the perturbation parameter ε (x-axis)

in variance in L^1 for successive reductions of the perturbation parameter ε (shown in Fig. 16b), clearly shows convergence of variance as $\varepsilon \to 0$. Moreover, Fig. 17 clearly indicates that in the $\varepsilon \to 0$ limit, the limit variance is nonzero. Hence, this strongly

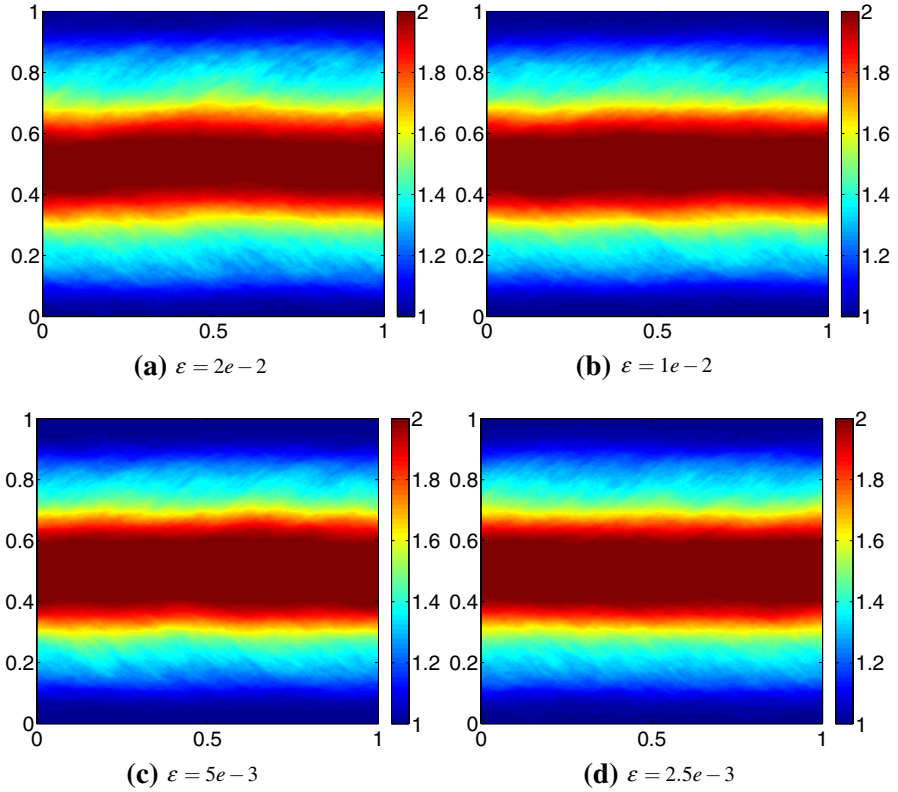


Fig. 15 Approximate sample means of the density for the Kelvin–Helmholtz problem (32) at time $t = 2$ and different values of perturbation parameter ε. All the computations are on a grid of 1024^2 mesh points and 400 Monte Carlo samples

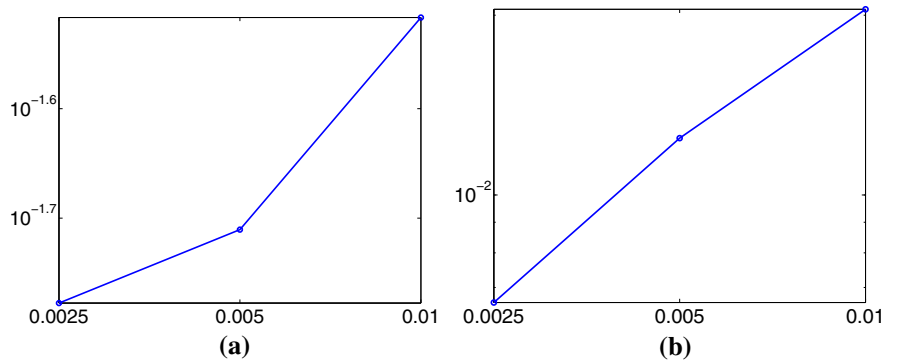


Fig. 16 Cauchy rates for the sample mean and the sample variance of the density (y-axis) for the Kelvin–Helmholtz problem (32) for different values of ε (x-axis). All the computations are on a grid of 1024^2 mesh points and 400 Monte Carlo samples. **a** Mean, **b** variance

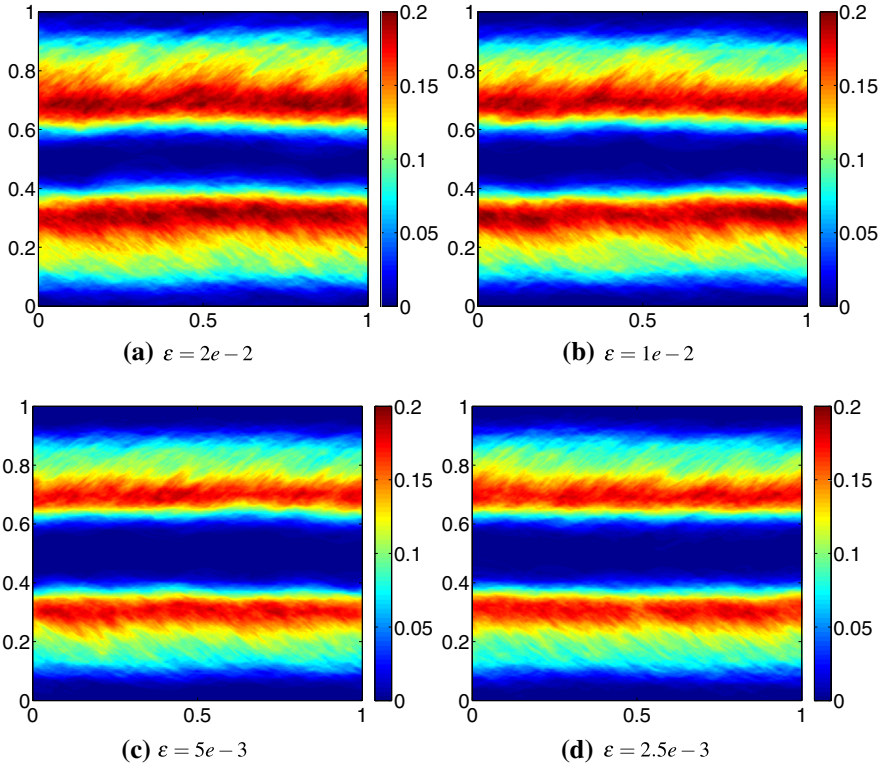


Fig. 17 Approximate sample variances of the density for the Kelvin–Helmholtz instability at time $t = 2$ and different values of perturbation parameter ε. All the computations are on a grid of 1024^2 mesh points and 400 Monte Carlo samples

suggests the fact that EMV solution can be *non-atomic, even for atomic initial data*. These results are consistent with the claims of Theorem 9.

To further demonstrate the non-atomicity of the resulting measure-valued solution, we have plotted the probability density functions (approximated by empirical histograms) for density at the points $x = (0.5, 0.7)$ and $x = (0.5, 0.8)$ in Fig. 18 for a fixed mesh of size 1024^2. We see that the initial unit mass centered at $\rho = 2$ ($\rho = 1$, respectively) at $t = 0$ is smeared out over time, and at $t = 2$ the mass has spread out over a range of values of ρ between 1 and 2.

Figure 19 shows the same quantities, but for a fixed time $t = 2$ over a series of meshes. Although a certain amount of noise seems to persist on the finer meshes—most likely due to the low number of Monte Carlo samples—it can be seen that the probability density functions seem to converge with mesh refinement.

6.3 Richtmeyer–Meshkov Problem

As a second numerical example, we consider the two-dimensional Euler equations (31) in the computational domain $x \in [0, 1]^2$ with initial data:

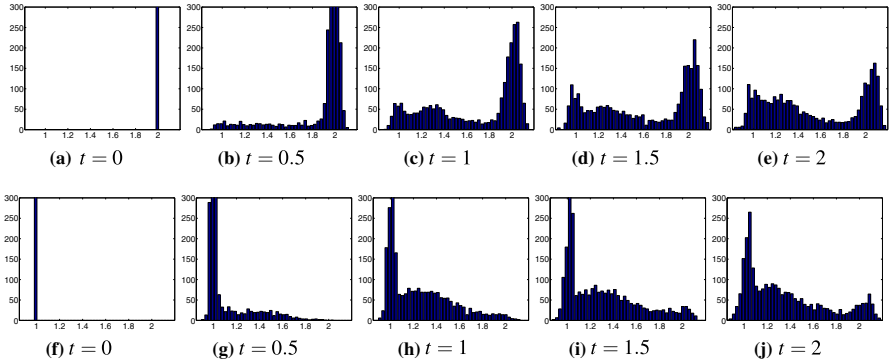


Fig. 18 Approximate PDF for density ρ at the points $x = (0.5, 0.7)$ (*first row*) and $x = (0.5, 0.8)$ (*second row*) on a grid of 1024^2 mesh points

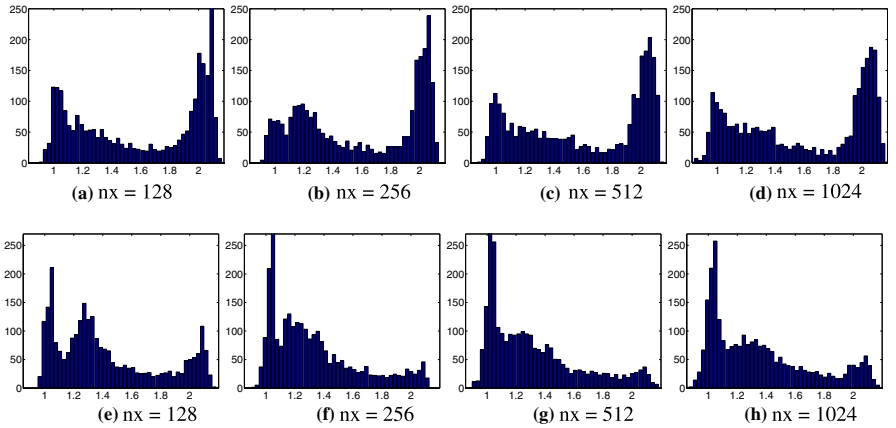


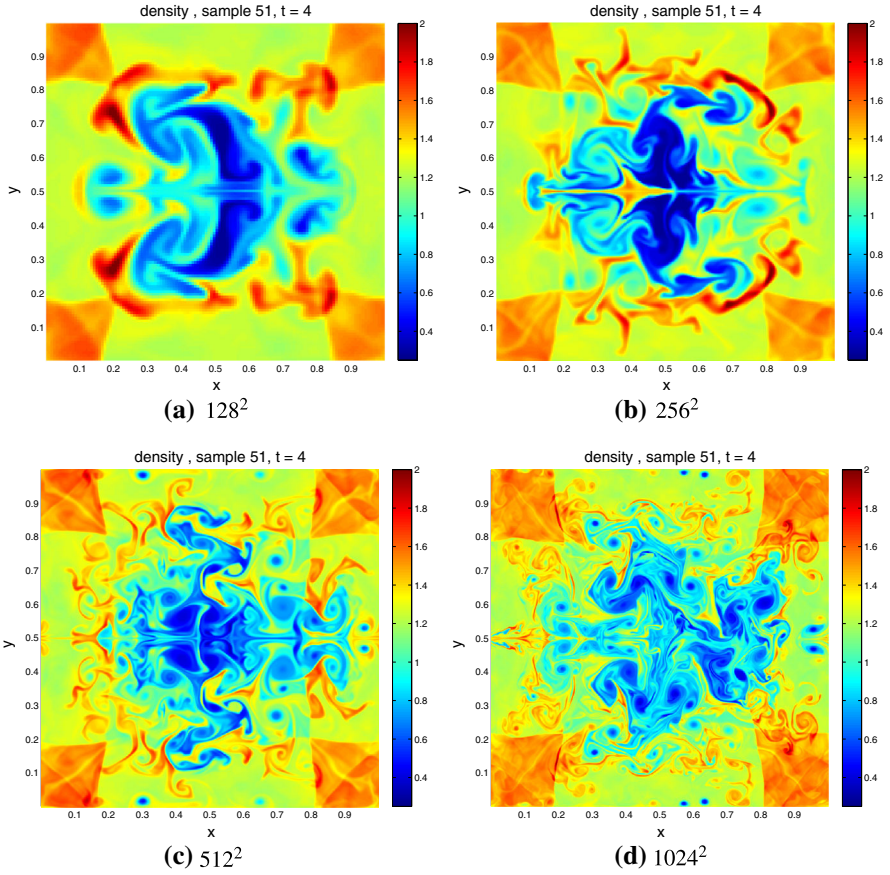
Fig. 19 Approximate PDF for density ρ at the points $x = (0.5, 0.7)$ (*first row*) and $x = (0.5, 0.8)$ (*second row*) on a series of meshes

$$p(x) = \begin{cases} 20 & \text{if } |x - (0.5, 0.5)| < 0.1 \\ 1 & \text{otherwise,} \end{cases}$$
$$\rho(x) = \begin{cases} 2 & \text{if } |x - (0.5, 0.5)| < I(x, \omega) \\ 1 & \text{otherwise,} \end{cases} \qquad w^1 = w^2 = 0. \tag{36}$$

The radial density interface $I(x, \omega) = 0.25 + \varepsilon Y(\varphi(x), \omega)$ is perturbed with

$$Y(\varphi, \omega) = \sum_{n=1}^{m} a^n(\omega) \cos\left(\varphi + b^n(\omega)\right), \tag{37}$$

where $\varphi(x) = \arccos((x_1 - 1/2)/|x - (0.5, 0.5)|)$ and a_n, b_n, k are the same as in Sect. 6.1.



(a) 128^2 **(b)** 256^2 **(c)** 512^2 **(d)** 1024^2

Fig. 20 Approximate density for a single sample for the Richtmeyer–Meshkov problem (36) for different grid resolutions at time $t = 4$

We repeat that the computational domain is $[0, 1]^2$, and we use periodic boundary conditions in both directions.

6.3.1 Lack of Sample Convergence

As in the case of the Kelvin–Helmholtz problem, we test whether numerical approximations for a single sample converge as the mesh is refined. To this end, we compute the approximations of the two-dimensional Euler equations with initial data (36) using a second-order MUSCL type finite volume scheme, based on the HLLC solver, and implemented in the FISH code [44]. The numerical results, presented in Fig. 20, show the effect of grid refinement on the density for a single sample at time $t = 4$. Note that by this time, the leading shock wave has exited the domain but has reentered from the corners on account of the periodic boundary conditions. Furthermore, this reentry shock wave interacts and strongly perturbs the interface forming a very complex

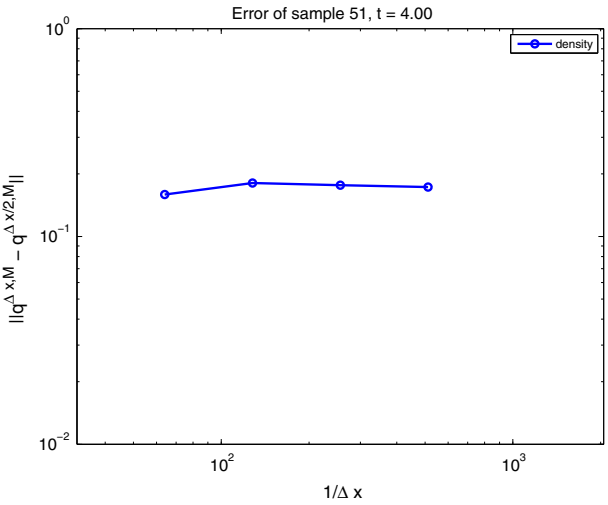


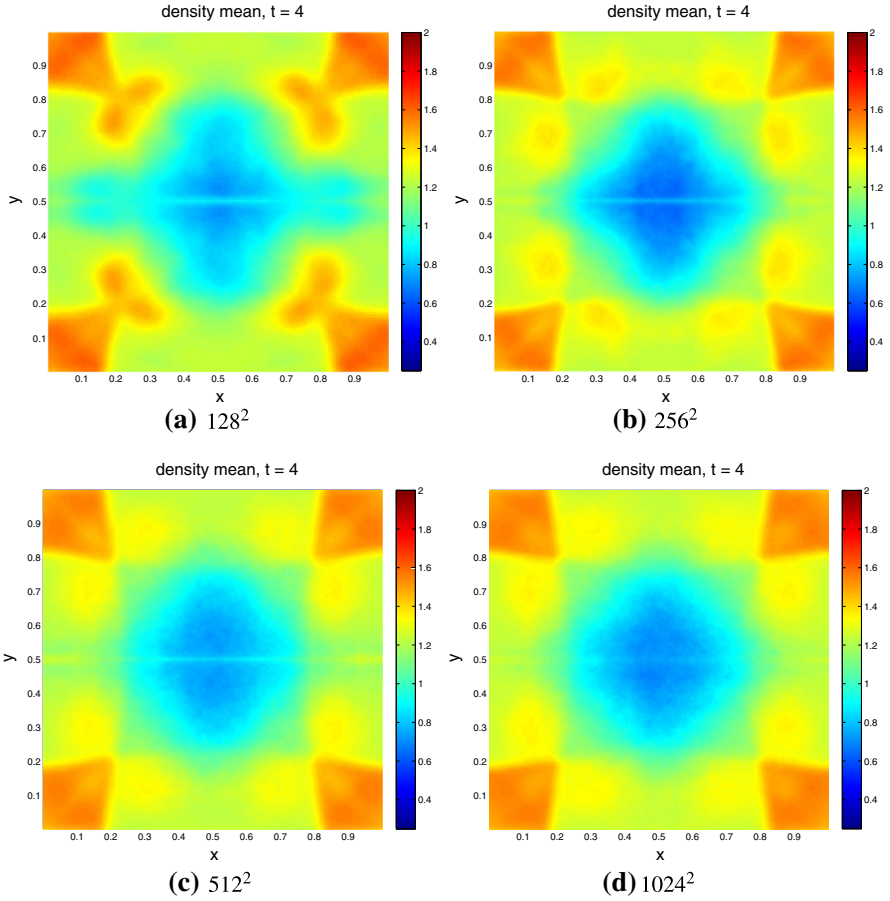
Fig. 21 Cauchy rates (7) for the density (y-axis) in a single sample of the Richtmeyer–Meshkov problem (36) at time $t = 4$, with respect to different grid resolutions (x-axis)

region of small-scale eddy-like structures. As seen from Fig. 20, there seems to be no convergence as the mesh is refined. This lack of convergence is quantified in Fig. 21, where we present differences in L^1 for successive mesh resolutions (7) and see that the approximate solutions for a single sample do not form a Cauchy sequence.

6.3.2 Convergence of the Mean and the Variance

Next, we test for convergence of statistical quantities of interest as the mesh is refined. First, we check the convergence of the mean through the Monte Carlo approximation (25) with $M = 400$ samples. The numerical results for the density at time $t = 4$ at different grid resolutions are presented in Fig. 22. The figure clearly shows that the mean converges as the mesh is refined. This convergence is further verified in Fig. 23a where we plot the difference in mean (33) for successive resolutions. This figure proves that the mean of the approximations forms a Cauchy sequence and hence converges. Figure 22 shows that small-scale features are averaged out in the mean and only large-scale structures, such as the strong reentrant shocks (mark the periodic boundary conditions) and mixing regions, are retained through the averaging process.

Next, we check for the convergence of the variance for the Richtmeyer–Meshkov problem (36). The results, shown in Fig. 24 for time $t = 4$, at different mesh resolutions and with 400 Monte Carlo approximations, clearly indicate that the variance of the approximate Young measures converge as the mesh is refined. This is also verified from Fig. 23b where the difference in L^1 of the variances at successive mesh resolutions is plotted and shown to form a Cauchy sequence. Furthermore, Fig. 24 also demonstrates that the variance is concentrated at the shocks and even more so in the mixing layer, around the original interface.



(a) 128^2 **(b)** 256^2 **(c)** 512^2 **(d)** 1024^2

Fig. 22 Mean density for the Richtmeyer–Meshkov problem with initial data (36) for different grid resolutions at time $t = 4$. All results are obtained with 400 Monte Carlo samples

6.4 Measure-Valued (MV) Stability

The above experiments clearly illustrate that the numerical procedure proposed here does succeed in computing an EMV solution of the underlying systems of conservation laws (11). Are the computed solutions stable? As argued in Sect. 3, uniqueness (stability) of EMV solutions for a general measure-valued initial data is not necessarily true, even for scalar conservation laws. Moreover, the scalar case suggests that at most a weaker concept of stability, that of MV stability can be expected for EMV solutions (see Terminology 3). As stated before, MV stability amounts to stability with respect to perturbations of atomic initial data. We examine this weaker notion of stability through numerical experiments.

To this end, we consider the Kelvin–Helmholtz problem as our test bed and investigate stability with respect to the following perturbations:

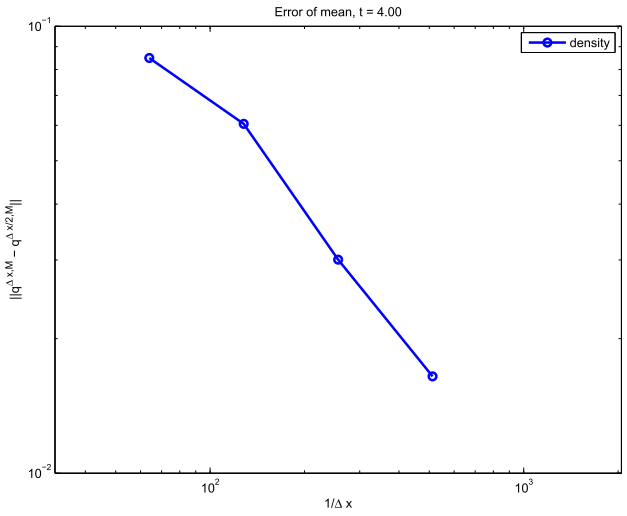
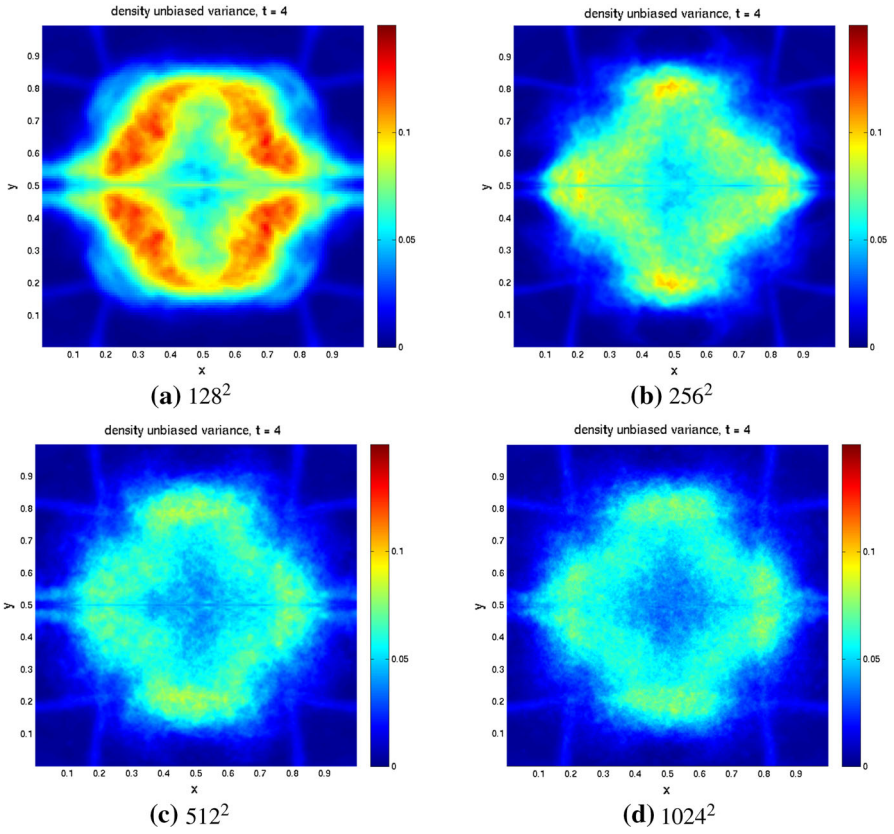


Fig. 23 Cauchy rates (33) for the mean and variance (y-axis) of the Richtmeyer–Meshkov problem (36) at time $t = 4$ and at different grid resolutions (x-axis). All results are obtained with 400 Monte Carlo samples. **a** Mean, **b** variance

6.4.1 Stability with Respect to Different Numerical Schemes

As a first check of MV stability, we consider the perturbed Kelvin–Helmholtz initial data (32) with a fixed perturbation size $\varepsilon = 0.01$ and compute approximate measure-valued solutions using Algorithm 5. Three different schemes are compared:

1. (Formally) second-order TeCNO2 scheme of [27].



(a) 128^2 **(b)** 256^2 **(c)** 512^2 **(d)** 1024^2

Fig. 24 Variance of the density with initial data (36) for different grid resolutions at time $t = 4$. All results are obtained with 400 Monte Carlo samples

2. Third-order TeCNO3 scheme of [27].
3. Second-order high-resolution finite volume scheme, based on the HLLC approximate Riemann solver, and implemented in the FISH code [44].

We will compare the mean and the variance of the approximate measures, at a resolution of 1024^2 points and 400 Monte Carlo samples, at time $t = 2$. As the mean and the variance with TeCNO2 scheme have already been depicted in Figs. 8d and 10d, respectively, we plot the mean and variance with the TeCNO3 and FISH schemes in Fig. 25. These results, together with the results for the TeCNO2 scheme (Figs. 8d, 10d), clearly show that mean and variance of the approximate measure-valued solution are very similar even though the underlying approximation schemes are different. In particular, comparing the TeCNO2 and TeCNO3 schemes, we remark that although both schemes have the same design philosophy (see [27] and Sect. 5), their formal order of accuracy is different. Hence, the underlying numerical viscosity operators are different. In spite of different numerical regularizations, both schemes seem to be converging to the same measure-valued solution—at least in terms of its first and second moments. This agreement is even more surprising for the FISH scheme of [44].

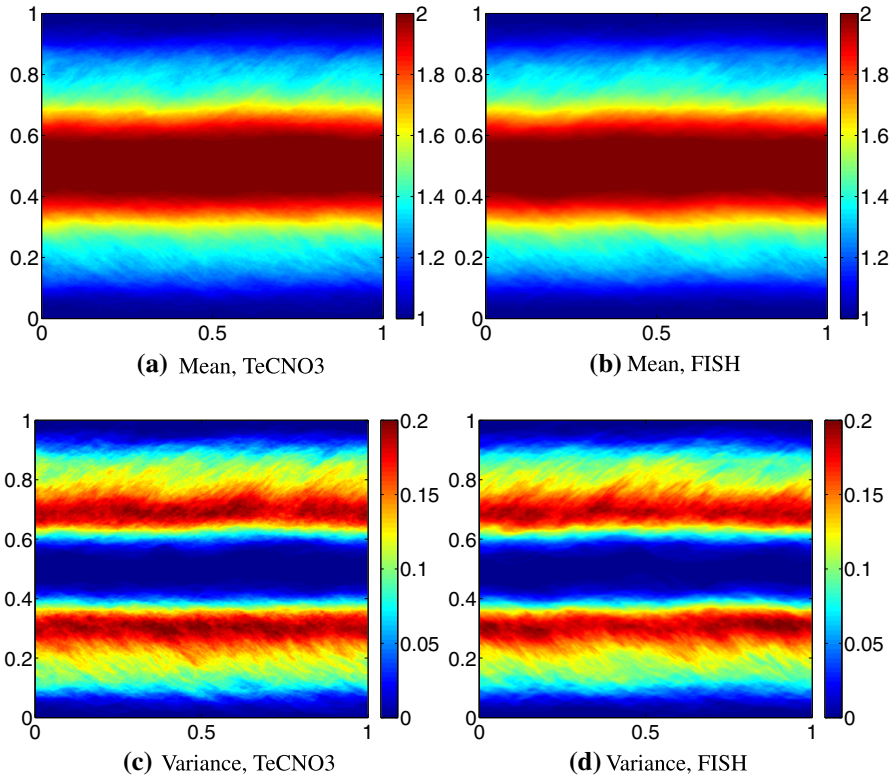


Fig. 25 Mean and variance of the density for the Kelvin–Helmholtz problem with initial data (32), at time $t = 2$ at a resolution of 1024^2 points and with 200 Monte Carlo samples. Different numerical schemes are compared. **a** Mean, TeCNO3; **b** mean, FISH; **c** variance, TeCNO3; **d** variance, FISH

This scheme utilizes a very different design philosophy based on HLLC approximate Riemann solvers and an MC slope limiter. Furthermore, it is unclear whether this particular scheme satisfies the discrete entropy inequality (22) or the weak BV bound (21). Nevertheless, the measure-valued solutions computed by this scheme seem to converge to the same EMV solution as computed by the TeCNO schemes. We have observed similar agreement between different schemes for smaller values of the perturbation parameter ε as well as in the Richtmeyer–Meshkov problem. Furthermore, all the three schemes agree with respect to higher moments as well. These numerical results at least indicate MV stability with respect to different numerical discretizations.

6.4.2 MV Stability with Respect to Different Perturbations

A more stringent test of MV stability is with respect to different types of initial perturbations. To be more specific, we consider the Kelvin–Helmholtz problem with the *phase* perturbations of (32) and compare them with *amplitude* perturbations (5) and (8). Note that for small values of the perturbation parameter ε, both the amplitude and phase perturbations are close to the atomic initial data (8) and to one another (for

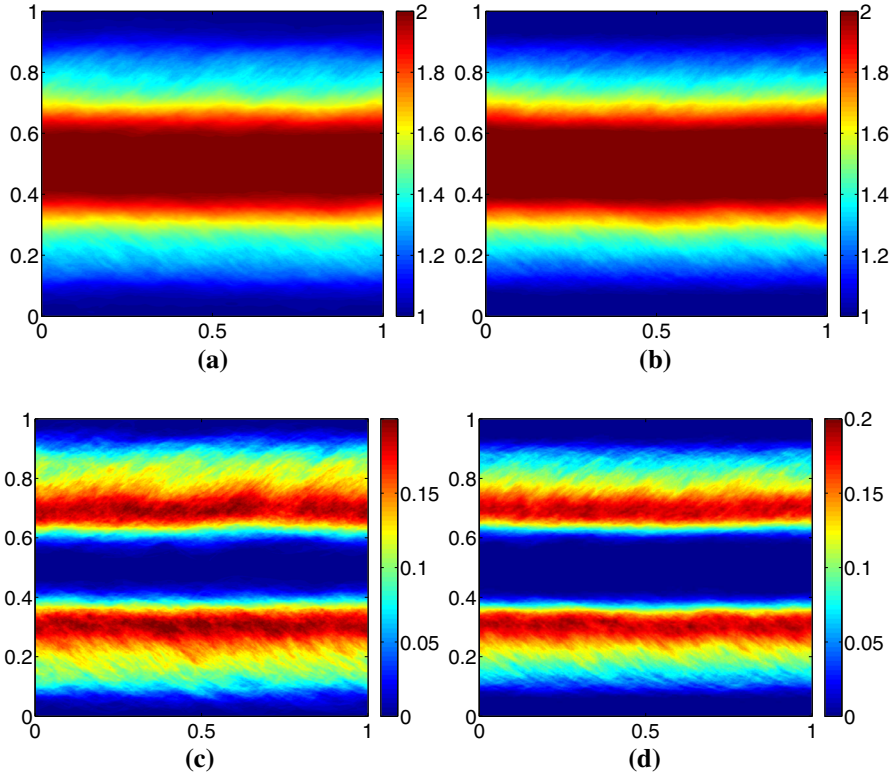


Fig. 26 Mean (*top*) and variance (*bottom*) of the density for the Kelvin–Helmholtz problem with different initial data: phase perturbations (32) (*left*) and amplitude perturbations (8), (5) (*right*), at time $t = 2$ at a resolution of 1024^2 points and with 400 Monte Carlo samples. All computations are with the TeCNO3 scheme. **a** Mean, phase perturbation; **b** mean, amplitude perturbation; **c** variance, phase perturbation; **d** variance, amplitude perturbation

instance in the Wasserstein metric). We test whether the resulting approximate MV solutions are also close. To this end, we compute the approximate measure-valued solutions with the phase perturbation and amplitude perturbation, for $\varepsilon = 0.0005$, with the TeCNO3 scheme, at a grid resolution of 1024^2 points and 400 Monte Carlo samples, and plot the results in Fig. 26. The results show that the mean and variance with different initial perturbations are very similar when the amplitude ε of the perturbations is small.

An even more stringent test of stability is provided by the following *phase* perturbation of the Kelvin–Helmholtz problem (32). The same set up (computational domain of $[0, 1]^2$ and periodic boundary conditions) as in the description of (32) is used but with an interface perturbation of the form:

$$I_j = I_j(x_1, \omega) := J_j + \varepsilon Y_j(x_1, \omega). \tag{38}$$

As in (32), we set $J_1 = 0.25$ and $J_2 = 0.75$ but with an interface variation of the form:

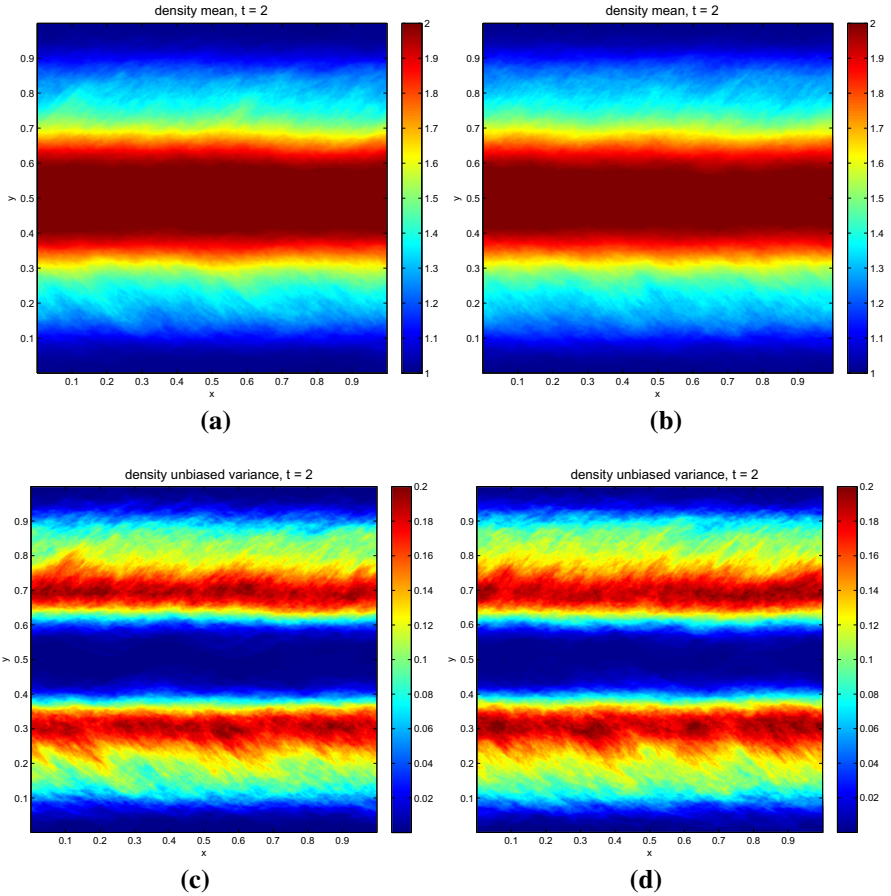


Fig. 27 Mean (*top*) and variance (*bottom*) of the density for the Kelvin–Helmholtz problem with different initial data: discontinuous phase perturbations (38) with uniformly distributed coefficients in (39) (*left*) and with (standard) normally distributed coefficients in (39) (*right*), at time $t = 2$ at a resolution of 1024^2 points and with 400 Monte Carlo samples. All computations are with the second-order FISH scheme. **a** Mean, uniform distribution; **b** mean, normal distribution; **c** variance, uniform distribution; **d** variance, normal distribution

$$Y_j(x_1, \omega) = \sum_{n=1}^{k} a_j^n \mathbb{1}_{A_n}, \quad j = 1, 2. \tag{39}$$

Here, $a_j^n = a_j^n(\omega) \in [-1/2, 1/2]$ are randomly chosen numbers from a uniform distribution. As a second variant, the a_j^n are drawn from the standard normal distribution. The A_n are equally spaced intervals, i.e., $A_n = [(n-1)h, nh)$ with $h = 1/32$. Thus, the initial interface perturbation is discontinuous, with uncorrelated random variation of the interface inside each interval. Such types of random initial data are motivated from observed or measured data, see [54]. A representative realization of this initial datum is shown in Fig. 5 (right).

The resulting approximate MV solutions, computed with a perturbation of size $\varepsilon = 0.005$, at time $t = 2$ and at a resolution of 1024^2 are shown in Fig. 27. The mean (top) and variance (bottom) are plotted. Results with the coefficients a_j^n, chosen from both an uniform distribution (left) and a standard normal distribution (right) are shown. As seen from the figure, the computed mean appears identical for the two choices of distributions. The same holds for the variance, where the resulting variances for both sets of distributions are very similar. Furthermore, they are also very similar to the corresponding statistical quantities, computed with the amplitude perturbation (5) and (8) as well as the sinusoidal phase perturbation (32) (compare with Fig. 26). Thus, we observe that the computed MV solutions are very similar to each other, even for four different sets of initial perturbations. Similar results were also observed for higher moments. This clearly indicates MV stability of the computed MV solution with the Kelvin–Helmholtz initial data.

7 Discussion

We conclude with a brief discussion on the highlights of the current paper which are put in perspective for future results. Currently, the notion of entropy solutions is the generic framework for interpreting the notion of solutions for $N \times N$ systems of hyperbolic conservation laws (1) in d spatial dimensions. Entropy solutions are *bounded functions* which satisfy the Eq. (2) and its associated entropy inequality(-ies) (3) in the sense of distributions. Though the existence and uniqueness of entropy solutions has been established for scalar conservation laws ($N = 1$) and for one-dimensional systems ($d = 1$), there are no known global existence and uniqueness (stability) results for generic multi-dimensional systems, when $N, d > 1$. In fact, recent papers [17–19] provide examples of multi-dimensional systems with infinitely many entropy solutions.

7.1 What Do the Numerical Experiments Tell Us

Despite a wide variety of numerical methods, such as finite volume, finite difference and discontinuous Galerkin methods that have been developed and successfully employed to approximate systems of conservation laws, none of these methods has been shown to converge to an entropy solution for a generic system of conservation laws. Given this background, we investigate here the issues of convergence of numerical approximations as well as the stability of the underlying entropy solutions. Our numerical experiments demonstrate that even state-of-the-art numerical methods may not necessarily converge as the mesh is refined. As shown in Figs. 3 and 6, finer and finer structures emerge as the grid is refined. The production of oscillations at finer and finer scales prevents convergence under mesh refinement. We also present numerical experiments that demonstrate the lack of stability of entropy solutions with respect to perturbations of initial data, see Figs. 4b and 13.

This lack of convergence to entropy solutions should not be considered as a failure of the numerical methods. Rather, they illustrate the shortcomings of the notion of entropy solutions to multi-dimensional systems of conservation laws. In particu-

lar, these experiments support the contention that entropy solutions are inadequate in describing some of the complex flow features that are modeled by systems of conservation laws such as the persistence of structures on finer and finer scales. Together with the recent results on the non-uniqueness of entropy solutions, our numerical evidence motivated us to seek a different, more versatile notion of solutions for these equations.

To this end, we focus on the notion of *entropy measure-valued (EMV) solutions*, first introduced by DiPerna in [22], see also [23]. We propose a measure-valued Cauchy problem (11) and seek solutions that are Young measures (parametrized probability measures). These *entropy measure-valued solutions* are sought to be consistent with the underlying equations in the sense of distributions (12) and satisfy a suitable form of the entropy inequality (13). *The main aim* of the current paper was then to design numerical procedure that can be rigorously shown to converge to an EMV solution. We work with an equivalent representation of measures as probability laws of random fields. The resulting initial random field is then evolved using a "reliable" entropy stable numerical scheme. The law of the resulting (random) weak* convergent approximations provides an approximation to the measure-valued solution. For a numerical scheme to be weak* convergent, it is required to satisfy a set of minimal criteria outlined in Theorem 6:

- *Uniform boundedness* of the approximations in L^∞;
- *Discrete entropy inequality*;
- *Space–time weak BV bound*.

The TeCNO schemes of [27] and the space–time DG schemes of [40] are examples of (formally) high-order schemes satisfying the discrete entropy inequality and weak BV bounds. The uniform L^∞ bound is a technical assumption that will be relaxed in a forthcoming paper [29]. Thus, we provide sufficient conditions that can guide the design of such "reliable" numerical methods for systems of conservation laws, with particular attention to multi-dimensional systems. Note that for systems of conservation laws, the above conditions play a role similar to that played by the well-known criteria of discrete maximum principle(s), entropy inequalities and the TVD property in the numerical analysis of scalar conservation laws.

The convergence of numerical approximations to an EMV solution of (11) is interpreted in the *weak** sense, namely that *statistics* of space–time averages of the unknowns converge as the mesh is refined. A Monte Carlo method is used to approximate the EMV solution, and we also show convergence of the resulting numerical procedure. To our knowledge, this *provides the first set of rigorous convergence results for numerical approximations of generic multi-dimensional systems of conservation laws*. These convergence results are illustrated by a large number of numerical experiments, and we make the following key observations:

- In general, there is no observed convergence of numerical approximations (neither in L^1 or in weaker norms) for single realizations (samples), with respect to increasing mesh resolutions. This has been demonstrated with two examples for the two-dimensional Euler equations.
- However, as predicted by the theory, statistical quantities of interest such as the mean and the variance (or even higher moments) of an ensemble of solutions do converge as the mesh is refined.

– In fact, a stronger convergence is observed. The approximate Young measures seem to converge in the strong sense (10) to an EMV solution.

The numerical approximation procedure, presented in Algorithm 8, was also employed to compute EMV solutions with respect to atomic initial data. In general, the computed measure-valued solution is *not necessarily atomic*, see Figs. 17, 18 and 19. This is particularly striking in the specific case of the Kelvin–Helmholtz problem, where an entropy solution (the steady-state data (8)), interpreted as an atomic entropy measure-valued solution, exists but is unstable.

This key observation implies that the solution operator can act to spread the support of the initial atomic measure. This bursting out of the initial atomic measure is, in our opinion, very significant. Just as the formation of shock waves precludes the existence of global classical solutions, leading to the replacement of point values with local averages as the appropriate solution concept, this observed bursting out of an initial atomic measure into a non-atomic measure implies that we have to look beyond integrable functions in order to obtain existence of solutions to systems of conservation laws. The concept of entropy measure-valued solutions, based on one-point statistics, appears to be a natural extension. In particular, given the proposed Algorithm 8, we are also able to address Lax's question raised in Sect. 4.2: What we are computing are the *statistics*—ensemble average, variance, etc.—of an entropy measure-valued solution.

7.1.1 Stability

As the results of this paper and the forthcoming paper [29] show, the convergence of numerical approximations also provides a (constructive) proof of existence for EMV solutions of (11). The questions of uniqueness and stability are much more delicate. From Remark 1, Example 1 and the results of [58] and references therein, we know that EMV solutions may not be unique if the initial measure is non-atomic, even for scalar conservation laws. We propose a weaker stability concept, that of measure-valued stability. This concept implies possible stability for the statistics of space–time averages in problems where the initial measure is close to atomic. Numerical experiments examining this weaker concept of stability were presented in Sect. 6.4. From these experiments, we observed that

– Different numerical schemes appear to converge to the same EMV solution as the mesh is refined.
– Different types of perturbations of atomic initial data were considered, and the resulting approximate EMV solutions seemed to converge to the same EMV solution, corresponding to atomic initial data.

These experiments indicate that our approximation procedure is indeed stable. Furthermore, they also suggest that the weaker notion of MV stability might be an appropriate framework to discuss the question of stability of EMV solutions.

Issues for Future Investigation

Our results raise several issues which are left open. We conclude this section with a few comments suggesting possible paths for future investigation.
On the notion of stability The only rigorous results available are of the measure-valued-strong uniqueness type, see Theorem 4 and [8,20]. Here, the stability is ensured when a classical solution (an atomic measure concentrated on a Lipschitz solution) is present. This also implies local (in time) uniqueness of EMV solutions for atomic initial data, concentrated on smooth functions. Given the paucity of rigorous stability results, there is a considerable open territory for future theoretical investigation of weaker concepts of stability, such as measure-valued stability for systems of conservation laws. Moreover, additional admissibility criteria such as entropy rate criteria of [16] or other variants might be necessary to ensure even MV stability of the EMV solution. This issue is dealt extensively in a forthcoming paper [25] where the concept of measure-valued solutions is further augmented with additional admissibility criteria, in the form of conditions on multi-point correlations, that increase the chance of singling out a unique solution.
Weak convergent schemes* As mentioned before, we provided here a numerical procedure, as well as sufficient conditions on numerical schemes, such that the approximations converge to an EMV solution. Some examples of schemes satisfying these criteria were presented. These results will hopefully encourage the development of other kinds of numerical schemes, such as of the WENO, RKDG and spectral viscosity type, that satisfy the abstract criteria of this paper, and hence converge to measure-valued solutions of systems of conservation laws, even in several space dimensions.
Computing the measure-valued solutions requires evaluation of phase space integrals. Our proposal in this paper was to employ Monte Carlo sampling. This procedure can be very expensive computationally, on account of the slow convergence with respect to the number of samples. We foresee the design of more computationally efficient methods by adapting schemes such as multi-level Monte Carlo [52–54], stochastic collocation finite volume methods [55] and gPC-based stochastic Galerkin methods [21], which have recently been developed to deal with uncertainty quantification for systems of conservation laws. Such extensions are the subject of ongoing research.

Acknowledgements S.M. and R.K. were supported in part by ERC STG. No. 306279, SPARCCLE. E.T. was supported in part by NSF Grants DMS10-08397, RNMS11-07444 (KI-Net) and ONR Grant N00014-1512094. Many of the computations were performed at CSCS Lugano through Project s345. SM thanks Prof. Christoph Schwab (ETH Zurich) for several helpful comments and suggestions.

Appendix 1: Young Measures

We provide here a very short introduction to Young measures. The reader may wish to consult [7,30] on the theory of Radon measures and probability measures and [2,3] on the theory of Young measures.

Probability Measures

A.1.1 We denote by $\mathcal{M}(\mathbb{R}^N)$ the set of finite Radon measures on $\mathbb{R}^N$, which are inner regular Borel measures μ with finite total variation $|\mu|(\mathbb{R}^N)$. Let $C_0(\mathbb{R}^N)$ be the space of continuous real-valued functions on $\mathbb{R}^N$ which vanish at infinity, equipped with the supremum norm. Then it can be shown (see, e.g., [30, Section 7.3]) that $\mathcal{M}(\mathbb{R}^N)$ can be identified with the dual space of $C_0(\mathbb{R}^N)$ through the pairing $\langle \mu, g\rangle = \int_{\mathbb{R}^N} g(\xi)\,\mathrm{d}\mu(\xi)$. We do not distinguish between these two equivalent definitions of $\mathcal{M}$. By a slight abuse of notation, we shall sometimes write $\langle \mu, g(\xi)\rangle = \int_{\mathbb{R}^N} g(\xi)\,\mathrm{d}\mu(\xi)$. We will be particularly interested in the pairing $\langle \mu, \mathrm{id}\rangle = \int_{\mathbb{R}^N} \xi\,\mathrm{d}\mu(\xi)$ between μ and the identity function $\mathrm{id}(\xi) = \xi$.

A.1.2 The duality between $C_0(\mathbb{R}^N)$ and $\mathcal{M}(\mathbb{R}^N)$ induces a weak* topology on $\mathcal{M}(\mathbb{R}^N)$, that of *weak* convergence*. A sequence $\mu^n \in \mathcal{M}(\mathbb{R}^N)$ converges *weak** to $\mu \in \mathcal{M}(\mathbb{R}^N)$ provided $\langle \mu^n, g\rangle \to \langle \mu, g\rangle$ for all $g \in C_0(\mathbb{R}^N)$. (This is also called weak or vague convergence, see [7,30].)

A.1.3 The set of probability measures on $\mathbb{R}^N$ is the subset

$$\mathcal{P}(\mathbb{R}^N) := \left\{\mu \in \mathcal{M}(\mathbb{R}^N) \;:\; \mu \geqslant 0,\ \mu(\mathbb{R}^N) = 1\right\}.$$

Let $\mathcal{P}^p(\mathbb{R}^N) \subset \mathcal{P}(\mathbb{R}^N)$ for $p \in [1, \infty)$ denote the set of probability measures μ such that $\langle \mu, |\xi|^p\rangle < \infty$. For $\mu, \rho \in \mathcal{P}^p(\mathbb{R}^N)$ the *Wasserstein metric* W_p is defined as

$$W_p(\mu, \rho) := \inf\left\{\int_{\mathbb{R}^N \times \mathbb{R}^N} |\xi - \zeta|^p\,\mathrm{d}\pi(\xi, \zeta) \;:\; \pi \in \Pi(\mu, \rho)\right\}^{1/p},$$

where $\Pi(\mu, \rho)$ is the set of probability measures on $\mathbb{R}^N \times \mathbb{R}^N$ with marginals μ and ρ:

$$\Pi(\mu, \rho) := \Big\{\pi \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^N) \;:\; \pi(A \times \mathbb{R}^N) = \mu(A),\ \pi(\mathbb{R}^N \times A) = \rho(A)\ \forall \text{ Borel } A \subset \mathbb{R}^N\Big\}.$$

It can be shown that W_p for any p metrizes the topology of weak convergence on $\mathcal{P}^p(\mathbb{R}^N)$ (see [1, Proposition 7.1.5] or [64, Chapter 7]).

A.1.4 Let $\mu, \rho \in \mathcal{P}(\mathbb{R})$, and let $F, G : \mathbb{R} \to [0, 1]$ be their distribution functions,

$$F(x) := \mu((-\infty, x]), \quad G(y) := \rho((-\infty, y]).$$

Then it can be shown that

$$W_p(\mu, \rho) = \left(\int_0^1 \left|F^{-1}(s) - G^{-1}(s)\right|^p \,\mathrm{d}s\right)^{1/p},$$

see [64, p. 75]. This gives rise to an efficient algorithm for computing the Wasserstein distance between discrete probability distributions. Let $x_1, \dots, x_n$ and $y_1, \dots, y_n$ be random numbers drawn from the probability distributions μ and ρ, respectively, and define the discrete distributions $\mu_n := (\delta_{x_1} + \cdots + \delta_{x_n})/n$ and $\rho_n := (\delta_{y_1} + \cdots + \delta_{y_n})/n$. By the law of large numbers, we have $\mu_n \to \mu$ and $\rho_n \to \rho$ weak* as $n \to \infty$, almost surely. Moreover, their distribution functions are

$$F_n(x) = \frac{\#\{x_j \,:\, x_j \leqslant x\}}{n}, \quad G_n(y) = \frac{\#\{y_j \,:\, y_j \leqslant y\}}{n}.$$

Hence, *if the sequences x_j and y_j are sorted in increasing order*, then

$$W_p(\mu_n, \rho_n)^p = \int_0^1 \left| F_n^{-1}(s) - G_n^{-1}(s) \right|^p \, \mathrm{d}s = \frac{1}{n} \sum_{j=1}^{n} |x_j - y_j|^p.$$

The latter expression is very easy to implement on a computer.

The analogous problem when $\mu, \rho \in \mathcal{P}(\mathbb{R}^N)$ is more complex, but can be solved in $O(n^3)$ time using the so-called Hungarian algorithm, see [56].

Young Measures

A.2.1 A *Young measure* from $D \subset \mathbb{R}^k$ to $\mathbb{R}^N$ is a function which maps $z \in D$ to a probability measure on $\mathbb{R}^N$. More precisely, a Young measure is a weak* measurable map $\nu : D \to \mathcal{P}(\mathbb{R}^N)$, that is, the mapping $z \mapsto \langle \nu(z), g \rangle$ is Borel measurable for every $g \in C_0(\mathbb{R}^N)$. We denote the image of $z \in D$ under ν by $\nu_z := \nu(z) \in \mathcal{P}(\mathbb{R}^N)$. The set of all Young measures from D into $\mathbb{R}^N$ is denoted by $\mathbf{Y}(D, \mathbb{R}^N)$. When $N = 1$ we write $\mathbf{Y}(D) := \mathbf{Y}(D, \mathbb{R})$.

A.2.2 A Young measure $\nu \in \mathbf{Y}(D, \mathbb{R}^N)$ is *uniformly bounded* if there is a compact set $K \subset \mathbb{R}^N$ such that $\operatorname{supp} \nu_z \subset K$ for all $z \in D$. Note that if ν is atomic, $\nu = \delta_u$, then ν is uniformly bounded if and only if $\|u\|_{L^\infty(D)} < \infty$.

A.2.3 If $u : \mathbb{R}^k \to \mathbb{R}^N$ is any measurable function, then $\nu_z := \delta_{u(z)}$ defines a Young measure, and we have $u(z) = \langle \nu_z, \mathrm{id} \rangle$ for every z. Conversely, we will say that a given Young measure ν is *atomic* if it can be written as $\nu = \delta_u$ for a measurable function u.

A.2.4 Two topologies on $\mathbf{Y}(D, \mathbb{R}^N)$ arise naturally in the study of Young measures: those of weak* and strong convergence. A sequence $\nu^n \in \mathbf{Y}(D, \mathbb{R}^N)$ converges *weak** to $\nu \in \mathbf{Y}(D, \mathbb{R}^N)$ if $\langle \nu^n, g \rangle \overset{*}{\rightharpoonup} \langle \nu, g \rangle$ in $L^\infty(D)$ for all $g \in C_0(\mathbb{R}^N)$, that is,

$$\int_D \varphi(z) \left\langle \nu_z^n, g \right\rangle \, \mathrm{d}z \to \int_D \varphi(z) \, \langle \nu_z, g \rangle \, \mathrm{d}z \quad \forall\, \varphi \in L^1(D).$$

We say that $\nu^n \in \mathbf{Y}(D, \mathbb{R}^N)$ converges *strongly* to $\nu \in \mathbf{Y}(D, \mathbb{R}^N)$ if

$$\left\| W_p(\nu^n, \nu) \right\|_{L^p(D)} \to 0$$

for some $p \in [1, \infty)$. If ν is atomic, $\nu = \delta_u$ for some $u : D \to \mathbb{R}^N$, then $\nu^n \to \nu$ strongly if and only if

$$\int_D \int_{\mathbb{R}^N} |\xi - u(z)|^p \, \mathrm{d}\nu_z^n(\xi)\mathrm{d}z \to 0.$$

A.2.5 The *fundamental theorem of Young measures* was first introduced by Tartar for L^∞-bounded sequences [63] and then generalized by Schonbek [59] and Ball [3] for sequences of measurable functions. We provide a further generalization: Every sequence $\nu^n \in \mathbf{Y}(D, \mathbb{R}^N)$ which does not "leak mass at infinity" (condition (40)) has a weak* convergent subsequence:

Theorem 13 *Let $\nu^n \in \mathbf{Y}(D, \mathbb{R}^N)$ for $n \in \mathbb{N}$ be a sequence of Young measures. Then there exists a subsequence ν^m which converges weak* to a nonnegative measure-valued function $\nu : D \to \mathcal{M}_+(\mathbb{R}^N)$ in the sense that*

(i) $\left\langle \nu_z^m, g \right\rangle \overset{*}{\rightharpoonup} \langle \nu, g \rangle$ *in* $L^\infty(D)$ *for all* $g \in C_0(\mathbb{R}^N)$,

and moreover satisfies

(ii) $\|\nu_z\|_{\mathcal{M}(\mathbb{R}^N)} \leqslant 1$ *for a.e.* $z \in D$;

(iii) If $K \subset \mathbb{R}^N$ *is closed and* $\operatorname{supp} \nu_z^n \subset K$ *for a.e.* $z \in D$ *and* n *large, then* $\operatorname{supp} \nu_z \subset K$ *for a.e.* $z \in D$.

Suppose further that for every bounded, measurable $E \subset D$, there is a nonnegative $\kappa \in C(\mathbb{R}^N)$ with $\lim_{|\xi| \to \infty} \kappa(\xi) = \infty$ such that

$$\sup_n \int_E \left\langle \nu_z^n, \kappa \right\rangle \, dz < \infty. \tag{40}$$

Then

(iv) $\|\nu_z\|_{\mathcal{M}(\mathbb{R}^N)} = 1$ *for a.e.* $z \in D$,

whence $\nu \in \mathbf{Y}(D, \mathbb{R}^N)$.

Proof The proof is a generalization of Ball [3].

Denote by $L_w^\infty(D; \mathcal{M}(\mathbb{R}^N))$ the set of weak* measurable functions $\mu : D \to \mathcal{M}(\mathbb{R}^N)$, equipped with the norm

$$\|\mu\|_{\infty,\mathcal{M}} := \operatorname*{ess\,sup}_{z \in D} \|\mu_z\|_{\mathcal{M}}.$$

From the fact that $C_0(\mathbb{R}^N)$ is separable, it can be shown (see [24, Theorem 8.18.2]) that $L_w^\infty(D; \mathcal{M}(\mathbb{R}^N))$ is isometrically isomorphic to the dual of $L^1(D; C_0(\mathbb{R}^N))$. The sequence μ^n is bounded in $L_w^\infty(D; \mathcal{M}(\mathbb{R}^N))$ since $\|\mu^n\|_{\infty,\mathcal{M}} \equiv 1$, and hence there is a $\mu \in L_w^\infty(D; \mathcal{M}(\mathbb{R}^N))$ and a weak* convergent subsequence μ^m of μ^n such that $\langle \mu^m, \Psi \rangle_{\infty,\mathcal{M}} \to \langle \mu, \Psi \rangle_{\infty,\mathcal{M}}$, or equivalently,

$$\int_D \left\langle \mu_z^m, \Psi(z, \cdot) \right\rangle \mathrm{d}z \to \int_D \langle \mu_z, \Psi(z, \cdot) \rangle \, \mathrm{d}z \quad \text{as } m \to \infty$$

for all $\Psi \in L^1(D; C_0(\mathbb{R}^N))$. In particular, letting $\Psi(z,\xi) = \varphi(z)g(\xi)$ for $\varphi \in L^1(D)$ and $g \in C_0(\mathbb{R}^N)$, we obtain *(i)*. We claim that $\mu_z \geqslant 0$ for a.e. $z \in D$. If not, then there would be a nonnegative $\Psi \in L^1(D; C_0(\mathbb{R}^N))$ such that $\int_D \langle \mu_z, \Psi(z,\cdot)\rangle \, \mathrm{d}z < 0$. But then

$$0 > \int_D \langle \mu_z, \Psi(z,\cdot)\rangle \, \mathrm{d}z = \lim_{m\to\infty} \int_D \left\langle \mu_z^m, \Psi(z,\cdot)\right\rangle \mathrm{d}z \geqslant 0$$

(since $\mu_z^m \geqslant 0$ for all z), a contradiction.

(ii) follows from the weak* lower semi-continuity of the norm $\|\cdot\|_{\infty,\mathcal{M}}$. To see that *(iii)* holds, let $g \in C_0(\mathbb{R}^N)$ be such that $g\big|_K = 0$. Since $\mu^m \to K$ in measure, it follows that $\langle \mu^m, g\rangle \to 0$ in measure (that is, $|\{z \in D : |\left\langle \mu_z^m, g\right\rangle| > \delta\}| \to 0$ for all $\delta > 0$). Hence,

$$\int_D \varphi(z) \, \langle \mu_z, g\rangle \, \mathrm{d}z = \lim_m \int_D \varphi(z) \left\langle \mu_z^m, g\right\rangle \mathrm{d}z = 0$$

and therefore $\langle \mu_z, g\rangle = 0$ for a.e. $z \in D$. This is precisely *(ii)*.

Assume now that (40) holds. Fix a set $E \subset D$ of finite, nonzero Lebesgue measure $|E|$, and denote the average integral over E as $⨍_E = \frac{1}{|E|}\int_E$. For every $R > 0$, we define

$$\theta_R(\xi) = \begin{cases} 1 & \kappa(\xi) \leqslant R \\ 1 + R - \kappa(\xi) & R < \kappa(\xi) \leqslant R+1 \\ 0 & R+1 < \kappa(\xi). \end{cases}$$

Then $\theta_R \in C_0(\mathbb{R}^N)$, so

$$\lim_m ⨍_E \left\langle \mu_z^m, \theta_R\right\rangle \mathrm{d}z = ⨍_E \langle \mu_z, \theta_R\rangle \, \mathrm{d}z \leqslant ⨍_E \|\mu_z\|_{\mathbb{R}} \mathrm{d}z \leqslant 1,$$

the last inequality following from the fact that $\|\mu_z\|_{\mathbb{R}} \leqslant 1$ for all z. Conversely,

$$0 \leqslant ⨍_E \left(1 - \left\langle \mu_z^m, \theta_R\right\rangle\right) \, \mathrm{d}z = ⨍_E \left\langle \mu_z^m, 1 - \theta_R\right\rangle \, \mathrm{d}z \leqslant \frac{1}{R} ⨍_E \left\langle \mu_z^m, \kappa\right\rangle \, \mathrm{d}z,$$

so (40) gives

$$\begin{aligned} 1 &\leqslant \lim_{R\to\infty} \lim_m ⨍_E \left\langle \mu_z^m, \theta_R\right\rangle \mathrm{d}z + \lim_{R\to\infty} \sup_m \frac{1}{R} ⨍_E \left\langle \mu_z^m, \kappa\right\rangle \, \mathrm{d}z \\ &= \lim_{R\to\infty} ⨍_E \langle \mu_z, \theta_R\rangle \, \mathrm{d}z \\ &\leqslant ⨍_E \|\mu_z\|_{\mathcal{M}(\mathbb{R}^N)} \mathrm{d}z \leqslant 1, \end{aligned}$$

whence $⨍_E \|\mu_z\|_{\mathcal{M}(\mathbb{R}^N)} \, \mathrm{d}z = 1$. Since $E \subset D$ is arbitrary, *(iv)* follows. □

A.2.6 An important special case of (40) is when $\kappa(\xi) = |\xi|^p$ for $1 \leqslant p < \infty$, which translates to the L^p bound

$$\sup_n \int_D \left\langle \mu^n, |\xi|^p \right\rangle \, \mathrm{d}z < \infty.$$

The case $p = \infty$ translates to the support of ν_z^n lying in a compact set $K \subset \mathbb{R}^N$ for a.e. z and all n. Part *(iii)* of Theorem 13 then holds for all $g \in C(\mathbb{R}^N)$, and condition (40) is automatically satisfied for any such κ. The latter is the original form of the theorem given by Tartar [63].

Random Fields and Young Measures

A.3.1 If $(\Omega, \mathcal{F}, P)$ is a probability space, $D \subset \mathbb{R}^k$ is a Borel set and $u : \Omega \times D \to \mathbb{R}^N$ is a random field (i.e., a jointly measurable function), then we can define its law by

$$\nu_z(F) := P\left(u(z) \in F\right) = P\left(\{\omega \, : \, u(\omega, z) \in F\}\right) \tag{41a}$$

for Borel subsets $F \subset \mathbb{R}^N$ of phase space, or equivalently,

$$\langle \nu_z, g \rangle := \int_\Omega g(u(\omega, z)) \, \mathrm{d}P(\omega) \tag{41b}$$

for $g \in C_0(\mathbb{R}^N)$. This defines a Young measure:

Proposition 1 *If $u : \Omega \times D \to \mathbb{R}^N$ is jointly measurable, then (41) defines a Young measure from D to $\mathbb{R}^N$.*

Proof First of all, for fixed $z \in D$ the set $\left\{\omega : u(\omega, z) \in U\right\}$ is P-measurable for Borel sets U. Indeed, if $w(\omega) := u(\omega, z)$ denotes the z-section of the measurable function $(\omega, y) \mapsto u(\omega, y)$, then $\left\{\omega : u(\omega, z) \in U\right\} = w^{-1}(U)$ is measurable.

We need to show that the definition of ν is independent of the choice of mapping in the equivalence classes of mappings from $\Omega \times D \to \mathbb{R}^N$. Let $\hat{u}, \tilde{u} : \Omega \times D \to \mathbb{R}^N$ be two mappings such that $\hat{u}(\omega, z) = \tilde{u}(\omega, z)$ for $P \times \lambda$-a.e. (ω, z). We apply Tonelli's theorem to find that

$$0 = \int_{\Omega \times D} \mathbb{1}_{\{\hat{u} \neq \tilde{u}\}}(\omega, z) \, d(P \times \lambda)(\omega, z) = \int_D P(\{\hat{u}(z) \neq \tilde{u}(z)\}) \, \mathrm{d}z.$$

Hence, $P(\hat{u}(z) \neq \tilde{u}(z)) = 0$ for a.e. $z \in D$, so for every Borel set $U \subset \mathbb{R}^N$,

$$P\left(\hat{u}(z) \in U\right) = P\left(\tilde{u}(z) \in U\right)$$

for a.e. $z \in D$.

Finally, ν is weak* measurable since

$$\langle \nu_z, g\rangle = \int_{\mathbb{R}^N} g(\xi)\,\mathrm{d}\nu_z(\xi) = \int_{\Omega} g(u(\omega, z))\,\mathrm{d}P(\omega),$$

which is measurable in z for any $g \in C_0(\mathbb{R}^N)$. □

A.3.2 It is well known that every measure on $\mathbb{R}^N$ can be realized as the law of a random variable. Here we show that for every Young measure ν, there is always a random field with law ν.

Proposition 2 *For every Young measure $\nu \in \mathbf{Y}(D, \mathbb{R}^N)$, there exist a probability space $(\Omega, \mathcal{F}, P)$ and a Borel measurable function $u : \Omega \times D \to \mathbb{R}^N$ such that u has law ν, i.e., for all Borel sets E,*

$$\nu_z(E) = P(u(\omega, z) \in E).$$

In particular, we can choose $(\Omega, \mathcal{F}, P)$ to be the Borel σ-algebra on $\Omega = [0, 1)$ with Lebesgue measure.

Proof The method of proof is standard, see, e.g., [6, Theorem 5.3].

We assume that $N = 1$. The generalization to $N > 1$ is straightforward but tedious. For $n \in \mathbb{N}$ and $j \in \mathbb{Z}$, we set

$$F_n^j := \begin{cases} (-\infty, -2^n) & \text{if } j = -2^{2n} \\ \left[2^{-n}(j-1), 2^{-n} j\right) & \text{if } j = -2^{2n}+1, \ldots, 2^{2n} \\ [2^n, \infty) & \text{if } j = 2^{2n}+1 \\ \emptyset & \text{otherwise.} \end{cases}$$

Let $p_n^j(z) := \sum_{l \leqslant j} \nu_z(F_n^l)$. Note that $p_n^j : \mathbb{R} \to [0, 1]$ is measurable for all n, j, and that $0 \leqslant p_n^{-j} \leqslant \ldots \leqslant p_n^j = 1$ for j large enough. Choose any $\xi_n^j \in F_n^j$, and for $\omega \in \Omega := [0, 1)$, define

$$u_n(\omega, z) := \xi_n^j \quad \text{for } j \text{ such that } p_n^{j-1}(z) \leqslant \omega < p_n^j.$$

We claim that u_n is measurable on the product σ-algebra between $\mathcal{F}$ and the Borel σ-algebra on D. Each function u_n takes only finitely many values ξ_n^j, so it suffices to show that $u_n^{-1}(\{\xi_n^j\})$ is measurable for every ξ_n^j. Indeed,

$$
\begin{aligned}
u_n^{-1}(\{\xi_n^j\}) &= \left\{(\omega, z) \in \Omega \times D \ : \ p_n^j(z) \leqslant \omega < p_n^{j+1}(z)\right\} \\
&= \Big(\Omega \times D\Big) \cap \left\{(\omega, z) \in \mathbb{R} \times D \ : \ p_n^j(z) \leqslant \omega\right\} \\
&\quad \cap \left\{(\omega, z) \in \mathbb{R} \times D \ : \ \omega < p_n^{j+1}(z)\right\},
\end{aligned}
$$

the intersection between the epigraph of p_n^j and the hypograph of p_n^{j+1}, which are measurable by the measurability of the functions p_n^j and p_n^{j+1}.

Because the partition $\{F_m^j\}_{j\in\mathbb{Z}}$ is a refinement of $\{F_n^j\}_{j\in\mathbb{Z}}$ whenever $m > n$, it follows that $|u_n(\omega, z) - u_m(\omega, z)| < \operatorname{diam}(F_n^j) = 2^{-n}$ for any (ω, z) whenever m, n are large enough. Hence, u_n converges pointwise to some function $u : \Omega \times D \to \mathbb{R}$, which is measurable by the measurability of each u_n.

Finally, for every $g \in C_0(\mathbb{R})$ and almost every $z \in D$, we have by Lebesgue's dominated convergence theorem

$$
\begin{aligned}
\int_\Omega g(u(\omega, z))\, \mathrm{d}P(\omega) &= \lim_n \int_\Omega g(u_n(\omega, z))\, \mathrm{d}P(\omega) \\
&= \lim_n \sum_j \nu_z(F_n^j) g(\xi_n^j) = \int_{\mathbb{R}} g(\xi)\, \mathrm{d}\nu_z(\xi).
\end{aligned}
$$

Hence, $u(\cdot, z)$ has law ν_z. □

Appendix 2: Proof of Theorem 11

Proof For any random field $\zeta : \Omega \to L^1(\mathbb{R}^d \times \mathbb{R}_+) \cap L^\infty(\mathbb{R}^d \times \mathbb{R}_+)$ on $(\Omega, \mathcal{F}, P)$, we denote the expectation with respect to the probability measure P as

$$
\mathbb{E}(\zeta) := \int_\Omega \zeta(\omega) \mathrm{d}P(\omega).
$$

For $1 \leqslant k \leqslant M$, denote

$$
\begin{aligned}
G(\omega) &= \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \psi(x, t) g\left(u^{\Delta x}(\omega; x, t)\right) \mathrm{d}x\mathrm{d}t, \\
G_k(\omega) &= \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \psi(x, t) g\left(u^{\Delta x, k}(\omega; x, t)\right) \mathrm{d}x\mathrm{d}t.
\end{aligned} \tag{42}
$$

Henceforth, we suppress the ω-dependence of G and G_k for notational convenience. The $L^2(P)$ error in the approximation can be written as

$$\begin{aligned}
\mathbb{E}\left(\left(\mathbb{E}(G)-\frac{1}{M}\sum_{k=1}^{M}G_k\right)^2\right) &= \mathbb{E}\left(\frac{1}{M^2}\left(\sum_{k=1}^{M}\left(\mathbb{E}(G)-G_k\right)\right)^2\right),\\
&= \mathbb{E}\left(\frac{1}{M^2}\left(\sum_{k=1}^{M}\left(\mathbb{E}(G)-G_k\right)^2+2\sum_{k=1}^{M}\sum_{l\neq k}\left(\mathbb{E}(G)-G_k\right)\left(\mathbb{E}(G)-G_l\right)\right)\right)\\
&= \underbrace{\frac{1}{M^2}\sum_{k=1}^{M}\mathbb{E}\left(\left(\mathbb{E}(G)-G_k\right)^2\right)}_{=:\,T_1}+\frac{2}{M^2}\sum_{k=1}^{M}\sum_{l\neq k}\underbrace{\mathbb{E}\left(\left(\mathbb{E}(G)-G_k\right)\left(\mathbb{E}(G)-G_l\right)\right)}_{=:\,T_2^{kl}}.
\end{aligned}$$

As $u^{\Delta x,1},\dots,u^{\Delta x,M}$ are independent and identically distributed, it follows from the definition of G_k that $G_1,\dots,G_M$ are independent and identically distributed random variables. Hence, $\mathbb{E}(G_k)=\mathbb{E}(G)$ and $\mathbb{E}(G_kG_l)=\mathbb{E}(G_k)\mathbb{E}(G_l)$ for all k,l. Consequently, a simple calculation shows that $T_2^{kl}=0$ for all $1\leqslant k,l\leqslant M$ and $k\neq l$.

The fact that $G_1,\dots,G_M$ are independent and identically distributed yields

$$T_1=\frac{1}{M}\left(\mathbb{E}(G^2)-\mathbb{E}(G)^2\right).$$

Hence,

$$\begin{aligned}
\mathbb{E}\left(\left(\mathbb{E}(G)-\frac{1}{M}\sum_{k=1}^{M}G_k\right)^2\right) &= \frac{1}{M}\left(\mathbb{E}(G^2)-(\mathbb{E}(G))^2\right)\\
&\leqslant \frac{1}{M}\left\|g(u^{\Delta x})\right\|^2_{L^\infty(\Omega\times\mathbb{R}^d\times\mathbb{R}_+)}\|\psi\|^2_{L^1(\mathbb{R}^d\times\mathbb{R}_+)} && \text{(by definition (42))}\\
&\leqslant \frac{C}{M} && \text{(by assumption (19a)).}
\end{aligned}$$

In conclusion, the sample mean

$$\frac{1}{M}\sum_{k=1}^{M}\int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)g\left(u^{\Delta x,k}(x,t)\right)\,\mathrm{d}x\mathrm{d}t$$

converges to the corresponding ensemble average, $\int_{\mathbb{R}_+}\int_{\mathbb{R}^d}\psi(x,t)\left\langle\nu^{\Delta x}_{x,t},g\right\rangle\,\mathrm{d}x\mathrm{d}t$ in $L^2(\Omega;P)$, with a convergence rate of $\frac{1}{\sqrt{M}}$. Taking a subsequence $M'\to\infty$, the convergence also holds P-almost surely. □

Appendix 3: Time Continuity of Approximations

From the time integration procedure (17b), we can show that the approximate MV solutions are time continuous. Consequently, the initial data is attained in a certain sense, and moreover, it is meaningful to evaluate the MV solution at a specific time t.

We state the theorem without proof, since the results are straightforward generalizations of "deterministic" counterparts.

Theorem 14 *Let $\psi \in C_c^1(\mathbb{R})$ and assume that (19a) and (19b) are satisfied. Let $\nu^{\Delta x}$ be generated by Algorithm 5. Then the functions*

$$\Psi^{\Delta x}(t) := \int_{\mathbb{R}} \psi(x) \left\langle \nu^{\Delta x}_{(x,t)}, \mathrm{id} \right\rangle \, dx$$

and

$$\Psi(t) := \int_{\mathbb{R}} \psi(x) \left\langle \nu_{(x,t)}, \mathrm{id} \right\rangle \, dx$$

are Hölder continuous with exponent $\gamma := \frac{r-1}{r}$ and with constant independent of Δx, and $\Psi^{\Delta x}(t) \to \Psi(t)$ as $\Delta x \to 0$ for a.e. $t \in [0, T]$. Moreover,

$$\Psi(0) = \lim_{t \to 0} \Psi(t) = \int_{\mathbb{R}} \psi(x) \langle \sigma_x, \mathrm{id} \rangle \, dx.$$

References

1. L. Ambrosio, N. Gigli and G. Savaré. *Gradient Flows*. Birkhäuser Basel, 2005.
2. E. J. Balder. *Lectures on Young Measures.* Université de Paris-Dauphine, 1995.
3. J. Ball. A version of the fundamental theorem for Young measures. In *PDEs and Continuum Models of Phase Transitions* (M. Rascle, D. Serre and M. Slemrod, eds.), Lecture Notes in Physics, vol. 344, Springer, 1989. 207–215.
4. T. J. Barth. Numerical methods for gas-dynamics systems on unstructured meshes. In *An Introduction to Recent Developments in Theory and Numerics of Conservation Laws*, Lecture Notes in Computational Science and Engineering volume 5, Springer, Berlin. Eds: D. Kroner, M. Ohlberger, and Rohde, C., 1999, 195–285.
5. S. Bianchini and A. Bressan. Vanishing viscosity solutions of nonlinear hyperbolic systems. *Ann. of Math.* (2) 161 (2005), no. 1, 223–342.
6. P. Billingsley. *Probability and Measure* 3rd ed. John Wiley & Sons Inc., 1995.
7. P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, Inc, 2008.
8. Y. Brenier and C. De Lellis and L. Székelyhidi Jr. Weak-Strong Uniqueness for Measure-Valued Solutions. *Comm. Math. Phys.,* 305 (2), 2011, 351–361.
9. A. Bressan, G. Crasta and B. Piccoli. *Well-posedness of the Cauchy problem for $n \times n$ systems of conservation laws.* Memoirs of the AMS, 146 (694), 2000.
10. Central Station: high-resolution non-oscillatory central schemes for non-linear conservation laws and related problems, www.cscamm.umd.edu/centpack/publications/.
11. G. Q. Chen and J. Glimm. Kolmogorov's theory of turbulence and inviscid limit of the Navier-Stokes equations in $\mathbb{R}^3$. *Comm. Math. Phys.* 310 (1), 2012, 267–283.
12. B. Cockburn, F. Coquel and P. G. LeFloch. Convergence of the finite volume method for multidimensional conservation laws. *SIAM J. Numer. Anal.*, 32 (3), 1995, 687–705.

13. B. Cockburn, C. Johnson, C. -W. Shu and E. Tadmor. Advanced Numerical Approximation of Nonlinear Hyperbolic Equations. Lecture notes in Mathematics 1697, 1997 C.I.M.E. course in Cetraro, Italy, June 1997 (A. Quarteroni ed.), Springer Verlag 1998.
14. B. Cockburn and C-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comput.*, 52, 1989, 411–435.
15. M. G. Crandall and A. Majda. Monotone difference approximations for scalar conservation laws. *Math. Comput.* 34, 1980, 1–21.
16. C. Dafermos. *Hyperbolic conservation laws in continuum physics.* Springer, Berlin, 2000.
17. C. De Lellis, L. Székelyhidi Jr. The Euler equations as a differential inclusion. *Ann. of Math.* (2) 170 (2009), no. 3, 1417–1436.
18. E. Chiodaroli, C. De Lellis, O. Kreml. Global ill-posedness of the isentropic system of gas dynamics. Preprint, 2013.
19. C. DeLellis and L. Székelyhidi Jr. On the admissibility criteria for the weak solutions of Euler equations. *Arch. Rational Mech. Anal.* 195 (2010), 225-260.
20. S. Demoulini and D. M. A. Stuart and A. E. Tzavaras. Weak-strong uniqueness of dissipative measure-valued solutions for polyconvex elastodynamics. *Archives of Rational Mechanics and Analysis*, 205 (3), 2012, 927–961.
21. B. Depres, G. Poette and D. Lucor. Uncertainty quantification for systems of conservation laws. *J. Comput. Phys.* 228 (2009), no. 7, 2443–2467.
22. R. J. DiPerna. Measure valued solutions to conservation laws. *Arch. Rational Mech. Anal.,* 88(3), 1985, 223–270.
23. R. J. DiPerna and A. Majda. Oscillations and concentrations in weak solutions of the incompressible fluid equations. *Comm. Math. Phys.* 108 (4), 1987, 667–689.
24. R. E. Edwards. *Functional Analysis. Theory and Applications.* Holt, Rinehart and Winston, Inc. (1965).
25. U. S. Fjordholm, S. Lanthaler and S. Mishra. Statistical solutions of hyperbolic conservation laws I. Theory, *In preparation*, 2015.
26. U. S. Fjordholm. S. Mishra and E. Tadmor. ENO reconstruction and ENO interpolation are stable. *FoCM* 13 (2), 2013, 139–159.
27. U. S. Fjordholm, S. Mishra and E. Tadmor. Arbitrary order accurate essentially non-oscillatory entropy stable schemes for systems of conservation laws. *SIAM J. Num. Anal* 50 (2), 2012, 544–573.
28. U. S. Fjordholm. *High-order accurate entropy stable numerical schemes for hyperbolic conservation laws.* ETH Zürich dissertation Nr. 21025, 2013.
29. U. S. Fjordholm, S. Mishra and E. Tadmor. Computation of measure valued solutions. Acta Numerica. *In preparation*, 2016.
30. G. B. Folland. *Real Analysis.* John Wiley & Sons Inc, 1999.
31. F. Fuchs, A. McMurry, S. Mishra, N. H. Risebro and K. Waagan. Approximate Riemann solver based high-order finite volume schemes for the MHD equations in multi-dimensions. *Comm. Comput. Phys* 9, 2011, 324–362.
32. H. Frid and I-S. Liu. Oscillation waves in Riemann problems for phase transitions. *Quart. Appl. Math.* 56 (1), 1998, 115–135.
33. H. Frid and I-S. Liu. Oscillation waves in Riemann problems inside elliptic regions for conservation laws of mixed type. *Z. Angew. Math. Phys.* 46 (1995), no. 6, 913–931.
34. S. Gottlieb, C.-W. Shu and E. Tadmor. High order time discretizations with strong stability properties. *SIAM. Review* 43, 2001, 89–112.
35. J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.* 18 (4), 1965, 697-715.
36. J. Glimm, J. Grove and Y. Zhang, Numerical Calculation of Rayleigh-Taylor and Richtmyer-Meshkov Instabilities for Three Dimensional Axisymmetric flows in Cylindrical and Spherical Geometries. Los Alamos Laboratory, Report# LA-UR99-6796, 1999.
37. E. Godlewski and P.A. Raviart. *Hyperbolic Systems of Conservation Laws.* Mathematiques et Applications, Ellipses Publ., Paris (1991).
38. A. Harten. High resolution schemes for hyperbolic conservation laws *J. Comput. Phys.* 49, 1983, 357–393.
39. A. Harten, B. Engquist, S. Osher and S. R. Chakravarty. Uniformly high order accurate essentially non-oscillatory schemes, III. *J. Comput. Phys.* 71 (2), 1987, 231–303.

40. A. Hiltebrand and S. Mishra. Entropy stable shock capturing streamline diffusion space-time discontinuous Galerkin (DG) methods for systems of conservation laws. *Num. Math.,* . 126 (1), 2014, 103-151.
41. J. Jaffre, C. Johnson and A. Szepessy. Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *Math. Model. Meth. Appl. Sci.,* 5(3), 1995, 367–386.
42. G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. *J. Comput. Phys.*, 126(1), 1996, 202–228.
43. C. Johnson and A. Szepessy. On the convergence of a finite element method for a nonlinear hyperbolic conservation law. *Math. Comput.*, 49 (180), 1987, 427–444.
44. R. Käppeli, S. C. Whitehouse, S. Scheidegger, U.-L. Pen and M. Liebendörfer. FISH: A Three-dimensional Parallel Magnetohydrodynamics Code for Astrophysical Applications. *The Astrophysical Journal Supplement*, 2011, 195, 20.
45. S. N. Kruzkhov. irst order quasilinear equations in several independent variables. *USSR Math. Sbornik.,* 10 (2), 1970, 217–243.
46. N.N. Kuznetsov. Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation. *USSR Comput. Math. and Math. Phys.* 16 (1976), 105–119.
47. P. D. Lax. Hyperbolic systems of conservation laws II. *Comm. Pure Appl. Math.* 10 (4), 1957, 537-566.
48. P. D. Lax. Mathematics and Physics. *Bull. AMS* 45(1), 2007, 135-152.
49. P. G. LeFloch, J. M. Mercier and C. Rohde. Fully discrete entropy conservative schemes of arbitrary order. *SIAM J. Numer. Anal.,* 40 (5), 2002, 1968–1992.
50. R. J. LeVeque. Finite volume methods for hyperbolic problems. *Cambridge university press,* Cambridge, 2002.
51. H. Lim, Y. Yu, J. Glimm, X. L. Li and D. H. Sharp. Chaos, transport and mesh convergence for fluid mixing. *Act. Math. Appl. Sin.,* 24 (3), 2008, 355–368.
52. S. Mishra and C. Schwab. Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data. *Math. Comput.*, 81(180), 2012, 1979–2018.
53. S. Mishra, Ch. Schwab and J. Sukys. Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions. *J. Comput. Phys* 231 (8), 2012, 3365–3388.
54. S. Mishra, Ch. Schwab and J. Sukys. Monte Carlo and multi-level Monte Carlo finite volume methods for uncertainty quantification in nonlinear systems of balance laws. *Uncertainty Quantification in Computational Fluid Dynamics.* Lecture Notes in Computational Science and Engineering Volume 92, 2013, 225–294
55. S. Mishra, N.H. Risebro, Ch. Schwab and S. Tokareva Numerical solution of scalar conservation laws with random flux functions. *Research report 2012-35*, SAM ETH Zürich.
56. J. Munkres. Algorithms for the Assignment and Transportation Problems *Journal of the Society for Industrial and Applied Mathematics*, 5 (1), 1957, 32–38.
57. B. Perthame and E. Tadmor. A kinetic equation with kinetic entropy functions for scalar conservation laws. *Communications in Mathematical Physics* 136, 1991, 501-517.
58. S. Schochet. Examples of measure-valued solutions *Comm. Par. Diff. Eqns.* 14 (5), 1989, 545–575.
59. M. Schonbeck. Convergence of solutions to nonlinear dispersion equations. *Comm. Par. Diff. Eqns.* 7 (8), 1982, 959–1000.
60. E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws, I. *Math. Comp.* 49, 1987, 91–103.
61. E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Act. Numerica,*, 2003, 451-512.
62. E. Tadmor. Convergence of spectral methods for nonlinear conservation laws. *SIAM Journal on Numerical Analysis* 26 (1989), 30–44.
63. L. Tartar. Compensated compactness and applications to partial differential equations. *Nonlinear analysis and mechanics: Heriot-Watt Symposium*, Vol. IV, Pitman, 1979, 136–212.
64. C. Villani. *Topics in Optimal Transportation.* American Mathematical Society, Graduate Studies in Mathematics, Vol. 58 (2013)