A moment approach for entropy solutions to nonlinear hyperbolic PDEs[*]

Swann Marx[1], Tillmann Weisser[1], Didier Henrion[1,2] and Jean Bernard Lasserre[1,3]

June 27, 2018

Abstract

We propose to solve polynomial hyperbolic partial differential equations (PDEs) with convex optimization. This approach is based on a very weak notion of solution of the nonlinear equation, namely the measure-valued (mv) solution, satisfying a linear equation in the space of Borel measures. The aim of this paper is, first, to provide the conditions that ensure the equivalence between the two formulations and, second, to introduce a method which approximates the infinite-dimensional linear problem by a hierarchy of convex, finite-dimensional, semidefinite programming problems. This result is then illustrated on the celebrated Burgers equation. We also compare our results with an existing numerical scheme, namely the Godunov scheme.

Keywords: nonlinear partial differential equation, convex optimization, moments and positive polynomials.

1 Introduction

This paper is concerned with the numerical study of scalar *nonlinear hyperbolic conservation laws*, a partial differential equation (PDE) which model numerous physical phenomena such as fluid mechanics, traffic flow or nonlinear acoustics [6], [34]. The existence and uniqueness of solutions to the associated Cauchy problem crucially depends on the flux and the initial condition [23]. Even if the solution is unique, its numerical computation is still a challenge – in particular when the solution has a shock, i.e., a discontinuity. Existing schemes based on discretization such as [16] suffer from numerical dissipation: the shock is smoothened in the numerical solution and cannot be represented accurately. In fact, sometimes the exact location of the shock is of crucial interest for applications. Note however that some existing numerical schemes are able

---

[1]CNRS; LAAS; Université de Toulouse; 7 avenue du colonel Roche, F-31400 Toulouse; France

[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 4, CZ-16206 Prague, Czechia

[3]IMT; UPS; Université de Toulouse; 7 avenue du colonel Roche, F-31400 Toulouse; France

to capture shock in the case where the conservation laws under consideration are linear, see e.g. [9].

In contrast with existing methods, a distinguishing feature of the numerical scheme presented in this paper is to *not* rely on discretization; it computes the solution in a given time-space window globally. From such a solution, the location of the shock at a given time can be computed up to the limits of machine precision. In our opinion this is a major advantage when compared to other numerical methods.

Measure-Valued Solutions While PDEs are usually understood in a weak sense, DiPerna proposed an even weaker notion of solution, the so-called *measure-valued solutions* (mv solutions for short) [10], which are based on Young measures, or parametrized (i.e. time and/or space dependent) probability measures. Young measures have originally been introduced in the context of calculus of variations and optimal control, where the velocity or more generally the control is relaxed from being a function of time to being a time-dependent probability measure on the control space, see e.g. [12, Part III] for an overview. Similarly, DiPerna introduced mv solutions to conservation laws as measures on the solution space, now depending on time and space.

Naturally, every weak solution gives rise to a mv solution when identifying a solution $y(t, x)$ with the Young measure $\delta_{y(t,x)}(dy)$. We say then that the mv solution is *concentrated* on (the graph of) the solution. In this paper, we are focusing on a setup where both weak and mv solutions are unique (hence identical). In this case, both solutions coincide via the identification just mentioned. Note however that our approach also applies without any change to situations where the mv solution is not concentrated, e.g., because of an initial condition that is not concentrated either.

In order to ensure uniqueness we rely on the notion of *entropy solutions* which has been extended to entropy mv solutions. Entropy is a concept from thermodynamics that makes reference to the fact that differences in physical systems, e.g., the densities of particles in a room, tend to adjust to each other. It is well-known that the entropy solution of a scalar nonlinear hyperbolic conservation law is unique. For the generalized situation things are more involved. However under suitable assumptions on the initial condition entropy, uniqueness of mv solutions can be proved.

Recently there has been an increasing interest in numerical schemes to compute mv solutions for hyperbolic conservation laws [15, 13]. In comparison, our approach rather computes numerically the *moments* of the unique mv solution and it does not rely on a time-space discretization. In a second step, the time and space dependent solution $(t, x) \mapsto y(t, x)$ can be recovered from the computed moments of the mv solution.

Generalized Moment Problem The key idea underlying the approach is to consider mv solutions as solutions to a particular instance of the *Generalized Moment Problem* (GMP) which is an infinite-dimensional optimization problem on appropriate spaces of measures, and where both the cost and the (possibly countably many) constraints are *linear* in the moments of the respective measures. Lasserre [24] showed that the GMP can be approximated as closely as desired by solving a hierarchy of convex *semidefinite programs* (SDP) of increasing size, called moment relaxations, provided that the data of the GMP are semi-algebraic; that is, the measures are supported on basic semi algebraic compact sets (i.e. bounded sets defined by finitely many polynomial inequalities and equations) and the involved functions are polynomial or semi-

2

algebraic functions (i.e. functions whose graphs are semi-algebraic sets). The duals to these moment relaxations are linear problems on polynomial sums of squares (SOS). This hierarchy of SDP relaxations is called the *moment-SOS (sums of squares) hierarchy*. By now, many problems from different fields of mathematics, including optimal control of ordinary differential equations [25], have been reformulated as particular instances of the GMP and then approximated or solved by the moment-SOS hierarchy. This paper is in the line of these former contributions. That is, (i) the mv solutions are viewed (or formulated) as solutions of a particular instance of the GMP, and (ii) the moments of mv solutions are approximated as closely as desired by solving a moment-SOS hierarchy.

Any optimal solution of each semidefinite relaxation at step d in the hierarchy provides information about the mv solution in the form of a sequence of its (approximated) moments, up to degree d; the higher is d the better is the approximation of its moments. As we restrict to measures with compact support, they are fully characterized from knowledge of the complete sequence of their moments. Interestingly, it is worth noting that in [15] it was already pointed out that the statistical moments of mv solutions are precisely the quantities of interest.

Contribution To the best of our knowledge, this work seems to be the first contribution where nonlinear PDEs are addressed *without time-space domain discretization* and using convex optimization *with a proof of convergence.* An original early attempt to compute mv solutions of nonlinear wave equations with linear programming was reported in [33], also in the presence of controls. In [25], the authors apply to moment-SOS hierarchy to solve optimal control problems of ordinary differential equations, and it was shown in [31] that it provides a sequence of subsolutions converging in norm to the viscosity solution of the Hamilton-Jacobi-Bellmann PDE, a particular nonlinear hyperbolic equation. In [28], nonlinear PDEs are discretized into large-scale sparse polynomial optimization problems, in turn solved with the moment-SOS hierarchy. More recently, bounds on functionals of solutions were obtained with SOS polynomials for nonlinear PDEs arising in fluid dynamics in [3] and for the nonlinear Kuramoto-Sivashinsky PDE in [17]. These works, however, focus only on the dual SOS problems, and they provide bounds with no convergence guarantees. They do not exploit the primal formulation of the problem on moments, which we believe to be crucial for convergence analysis. In the recent work [14], the authors compute mv solutions for the equations of compressible and incompressible inviscid fluid dynamics, with the help of discretization algorithms based on Monte Carlo methods. Even more recently, in [1] the author has proposed a convex formulation for the cla,ssical solution to nonlinear hyperbolic PDEs and he proves that the entropy solution to the Burgers equation might be recovered also via this optimization problem. However, this paper does not provide a numerical scheme. In the concurrent work [22], the authors propose to use the moment-SOS hierarchy in a much more general setting of a controlled polynomial PDE. However, at that level of generality, there is no proof that the numerical scheme will converge to an appropriate solution of the PDE. For more references on previous attempts to use convex optimization for solving and controlling PDEs, the reader is referred to the introduction of [22].

Outline This paper is organized as follows. Section 2 introduces different notions of solutions for scalar conservation laws and provides some links between these notions. Section 3 introduces the Moment-SOS hierarchy, proves that the mv solution framework can be written as an instance of the GMP, and explains how one may interpret the moment solutions. The focus of Section 4 is on a numerical study of the Burgers equation. Finally, Section 5 collects some concluding remarks and further research issues to be addressed.

Notation: If $\mathcal{X}$ is a topological space, let $\mathscr{C}(\mathcal{X})$ resp. $\mathscr{C}_0(\mathcal{X})$ resp. $\mathscr{C}_c^1(\mathcal{X})$ denote the space of functions on $\mathcal{X}$ that are continuous resp. continuous and vanishing at infinity resp. continuously differentiable with compact support. For $p \geq 1$, the Lebesgue space $\mathscr{L}^p(\mathcal{X})$ consists of functions on $\mathcal{X}$ whose p-norms are bounded. The set of signed resp. positive Borel measures is denoted $\mathscr{M}(\mathcal{X})$ resp. $\mathscr{M}(\mathcal{X})_+$. The set of probability measures on $\mathcal{X}$ is denoted by $\mathscr{P}(\mathcal{X})$ and it consists of elements $\mu \in \mathscr{M}(\mathcal{X})_+$ such that $\mu(\mathcal{X}) = 1$. The measure $\lambda_{\mathcal{X}} \in \mathscr{P}(\mathcal{X})$ denotes the normalized Lebesgue measure on $\mathcal{X}$. Given a vector $\mathbf{w} = (w_1 \ \ldots \ w_n)$, we denote by $\mathbb{R}[\mathbf{w}]$ the ring of real multivariate polynomials in the variables $w_1, \ldots, w_n$.

2 Notions of solutions

We start with a brief overview of different notions of solutions to scalar polynomial PDEs. For details, we refer to [6] for weak solutions and [29] for measure-valued solutions. The aim of this section is to give a clear link between these two concepts of solutions.

2.1 Weak and entropy solutions

In order to study mv solutions, it is instructive to revisit the classical concept of weak solutions first. Consider therefore the Cauchy problem

$$\frac{\partial y}{\partial t}(t, x) + \frac{\partial f(y)}{\partial x}(t, x) = 0, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}, \tag{1a}$$

$$y(0, x) = y_0(x), \quad x \in \mathbb{R}, \tag{1b}$$

where (1a) is a scalar hyperbolic conservation law with $f \in \mathscr{C}^1(\mathbb{R})$ and (1b) provides an initial condition $y_0 \in \mathscr{L}^1(\mathbb{R}) \cap \mathscr{L}^\infty(\mathbb{R})$. Note that (1a) encompasses, among others, the well-known Burgers equation if one sets $f(y) = \frac{1}{2}y^2$.

Even if the initial condition y_0 is smooth, solutions to (1) might be discontinuous (see [11, p. 143] for the case of the Burgers equation). Solutions to this problem are hence usually understood in the following weak sense.

Definition 1 (Weak solution). *A function $y \in \mathscr{L}^\infty(\mathbb{R}_+ \times \mathbb{R})$ is a weak solution to (1) if, for all test functions $\psi_1 \in \mathscr{C}_c^1(\mathbb{R}_+ \times \mathbb{R})$, it satisfies*

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left(\frac{\partial \psi_1(t, x)}{\partial t} y(t, x) + \frac{\partial \psi_1(t, x)}{\partial x} f(y(t, x)) \right) dx \, dt + \int_{\mathbb{R}} \psi_1(0, x) y_0(x) \, dx = 0. \tag{2}$$

In general, weak solutions to (1) are not unique. However it can be shown (see e.g. [23]) that among all possible weak solutions, only one has a physical meaning. This solution is called an *entropy solution* and it can be characterized as follows.

Definition 2 (Entropy pair/entropy solution). *(i) A pair of functions $\eta, q \in \mathscr{C}^1(\mathbb{R})$ is called an* entropy pair *for (1a) if η is strictly convex and $q' = f'\eta'$.*

(ii) A weak solution $y \in \mathscr{L}^\infty(\mathbb{R}_+ \times \mathbb{R})$ of (1) is an entropy solution *if, for all entropy pairs and all non-negative test functions $\psi_2 \in \mathscr{C}^1_c(\mathbb{R}_+ \times \mathbb{R})$, it satisfies*

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left(\frac{\partial \psi_2}{\partial t} \eta(y) + \frac{\partial \psi_2}{\partial x} q(y) \right) dx\, dt + \int_{\mathbb{R}} \psi_2(0, x) \eta(y_0)(x)\, dx \geq 0. \qquad (3)$$

2.2 Measure-valued solutions

Generally, regularity results of conservation laws are obtained from regularized conservation laws

$$\frac{\partial}{\partial t} y^\varepsilon + \frac{\partial}{\partial x} f(y^\varepsilon) - \varepsilon \frac{\partial^2}{\partial x^2} y^\varepsilon = 0,$$

where $\varepsilon > 0$ is a fixed parameter. Then one studies the limit of solutions y^ε as ε goes to 0 and tries to retrieve some of regularity properties of the latter equation for the conservation law. However, on the one hand, regularized solution y^ε may or may not converge to a weak solution y of (1). This is due of a lack of reflexivity of the space $\mathscr{L}^\infty$. On the other hand, regularized solutions y^ε necessarily converge to a measure-valued (mv) solution. This notion builds upon the concept of a Young measure.

Definition 3 (Young measure). *A Young measure on a Euclidean space $\mathcal{X}$ is a map $\mu : \mathcal{X} \to \mathscr{P}(\mathbb{R})$, $\xi \mapsto \mu_\xi$, such that for all $g \in \mathscr{C}_0(\mathbb{R})$ the function $\xi \mapsto \int_{\mathbb{R}} g(y)\mu_\xi(dy)$ is measurable.*

Later, mv solutions have also proved to be useful in the study of problems more general than (1), where the initial condition (1b) is replaced by a Young measure parametrized in space (see e.g. [15] and the references therein). The generalized problem is to find a Young measure $\mu_{(t,x)}$ which satisfies the following Cauchy problem:

$$\partial_t \langle \mu_{(t,x)}, y \rangle + \partial_x \langle \mu_{(t,x)}, f(y) \rangle = 0, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}, \qquad (4a)$$

$$\mu_{(0,x)} = \sigma_0, \quad x \in \mathbb{R}, \qquad (4b)$$

where $\langle \cdot, \cdot \rangle$ denotes integration of a measure $\mu \in \mathscr{M}(\mathbb{R})$ against a function $g \in \mathscr{C}(\mathbb{R})$:

$$\langle \mu, g \rangle := \int_{\mathbb{R}} g(y)\mu(dy).$$

In (4) the measure σ_0 is a given Young measure on $\mathbb{R}$, and f is a continuously differentiable function on $\mathbb{R}$. The conservation law (4a) has to be understood in the sense of distributions, i.e.:

Definition 4 (Measure-valued solution). *A Young measure is a measure-valued (mv) solution to (4) if, for all test functions $\psi_1 \in \mathscr{C}^1_c(\mathbb{R}_+ \times \mathbb{R})$, it satisfies*

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left(\frac{\partial \psi_1(t, x)}{\partial t} \langle \mu_{(t,x)}, y \rangle + \frac{\partial \psi_1(t, x)}{\partial x} \langle \mu_{(t,x)}, f(y) \rangle \right) dx\, dt + \int_{\mathbb{R}} \psi_1(0, x) \langle \sigma_0, y \rangle dx = 0.$$
$$\qquad (5)$$

Note that the weak solution y has been replaced by a time-space parametrized probability measure μ supported on the range of y. Whereas a weak solution is requested to satisfy (2), only averages of the mv solution are considered in (5). It is easy to see that every weak solution induces a mv solution via the canonical embedding $y(t,x) \mapsto \delta_{y(t,x)}$. As in the case of weak solution, an entropy condition is needed in order to select solutions with a physical meaning. Quite in analogy to entropy solutions, entropy mv solutions are defined as follows.

Definition 5 (Entropy measure-valued solution)**.** *An mv solution μ is an entropy mv solution to (4) if, for all entropy pairs (η, q) and all non-negative test functions $\psi_2 \in \mathscr{C}_c^1(\mathbb{R}_+ \times \mathbb{R})$, it satisfies*

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left(\frac{\partial \psi_2(t,x)}{\partial t} \langle \mu_{(t,x)}, \eta \rangle + \frac{\partial \psi_2(t,x)}{\partial x} \langle \mu_{(t,x)}, q \rangle \right) dx \, dt + \int_{\mathbb{R}} \psi_2(0,x) \langle \sigma_0, \eta \rangle dx \geq 0.$$
(6)

Remark 1. *Again it is straightforward to see that entropy solutions are entropy mv solutions via the canonical embedding $y(t,x) \mapsto \delta_{y(t,x)}$. However, as demonstrated on an example in [15, p. 775], in contrast with entropy solutions, entropy mv solutions are not necessarily unique.*

We have seen that the concept of mv solutions is weaker than the concept of weak solutions. Hence mv solutions are a relaxation of weak solutions: every weak solution is also an mv solution, but the set of mv solutions can be larger than the set of weak solutions. However, the following result states that considering mv solutions is not a relaxation. To be more precise, when the initial measure in (4b) is concentrated on (the graph of) the initial condition in (1b), then the entropy mv solution to (4) is unique and concentrated on (the graph of) the (unique) entropy solution to (1).

Theorem 1 (Concentration of the entropy mv solution)**.** *Let C be the Lipschitz constant of the function f. Let y be an entropy solution and μ be an entropy mv solution to (1). Then, for all $T \geq 0$ and all $r \geq 0$, it holds*

$$\int_{|x| \leq r} \langle \mu_{(t,x)}, |y - y(T,x)| \rangle dx \leq \int_{|x| \leq r+CT} \langle \sigma_0, |y - y_0(x)| \rangle dx.$$
(7)

In particular, if $\sigma_0 = \delta_{y_0(x)}$, then $\mu_{(t,x)} = \delta_{y(t,x)}$ for all $t \in [0,T]$ and all x such that $|x| \leq r$.

Remark 2. *The proof of Theorem 1 is similar to the one provided in [23] and it is postponed to Appendix A. It is based on the doubling variable strategy, using the following family of entropy pairs:*

$$\eta_v(y) = |y - v|, \quad q_v(y) = \mathrm{sign}(y - v)(f(y) - f(v))$$
(8)

parametrized in $v \in \mathbb{R}$. In [26], it has been proved that a linear combination of these entropy pairs, together with the convex hull of linear functions, generate all entropy pairs. In other words, to prove Theorem 1 for every entropy pair it is enough to consider the entropy pairs (8).

Moreover, note that initially the doubling variable strategy has been used to prove uniqueness of the solution to scalar nonlinear conservations laws. The main drawback is that the entropy solution has to satisfy this inequality for all convex pairs. However for the specific case of the Burgers equation, it is shown in [7] and [30] that one may consider only one convex pair.

2.3   An emphasis on compact sets

In practice, one computes or approximates the solution on compact subsets, so let

$$\mathbf{T} := [0, T], \quad \mathbf{X} := [L, R] \tag{9}$$

be the respective domains of time t and space x, for fixed (but arbitrary) constants T, L, R. After scaling, we assume without loss of generality that $T = R - L = 1$.

Note that the entropy inequality induces a stability property:

$$\|y(t, \cdot)\|_{\mathscr{L}^\infty(\mathbf{X})} \leq \|y_0\|_{\mathscr{L}^\infty(\mathbb{R})}, \ \forall t \geq 0 \tag{10}$$

see e.g. [6, Theorem 6.2.4]. Since y_0 is bounded in $\mathscr{L}^\infty$, it follows from the maximum principle [6, Theorem 6.3.2] that $y(t, .)$ is bounded in $\mathscr{L}^\infty$ for all $t \geq 0$. Hence, we can consider that y takes values in the following compact set

$$\mathbf{Y} := [\underline{y}, \bar{y}], \tag{11}$$

where the bounds $\underline{y} := \mathrm{ess} \inf_{x \in \mathbb{R}} y_0(x)$ and $\bar{y} := \mathrm{ess} \sup_{x \in \mathbb{R}} y_0(x)$ depend on the initial condition. On $\mathbf{T} \times \mathbf{X}$, the polynomial hyperbolic equations given in (1) reads:

$$\begin{cases} \dfrac{\partial y}{\partial t} + \dfrac{\partial f(y)}{\partial x} = 0, & (t, x) \in \mathbf{T} \times \mathbf{X}, \\ y(0, x) = y_0(x), & x \in \mathbf{X}. \end{cases} \tag{12}$$

Definition 6 (Entropy solution on compact sets). *A weak solution y is an entropy solution to (12) if, for all test functions $\psi_1 \in \mathscr{C}^1(\mathbb{R}_+ \times \mathbb{R})$, it satisfies*

$$\int_{\mathbf{T}} \int_{\mathbf{X}} \left(\frac{\partial \psi_1}{\partial t} y + \frac{\partial \psi_1}{\partial x} f(y) \right) dx dt + \int_{\mathbf{X}} \psi_1(0, x) y_0(x) dx - \int_{\mathbf{X}} \psi_1(T, x) y(T, x) dx$$
$$+ \int_{\mathbf{T}} \psi_1(t, L) y(t, L) dt - \int_{\mathbf{T}} \psi_1(t, R) y(t, R) dt = 0 \tag{13}$$

and, for all convex pairs (p, q) and all non-negative test functions $\psi_2 \in \mathscr{C}^1(\mathbb{R}_+ \times \mathbb{R})$, it satisfies

$$\int_{\mathbf{T}} \int_{\mathbf{X}} \left(\frac{\partial \psi_2}{\partial t} \eta(y) + \frac{\partial \psi_2}{\partial x} q(y) \right) dx dt + \int_{\mathbf{X}} \psi_2(0, x) \eta(y_0)(x) dx - \int_{\mathbf{X}} \psi_2(T, x) \eta(y(T, x)) dx$$
$$+ \int_{\mathbf{T}} \psi_2(t, L) q(y(t, L)) dt - \int_{\mathbf{T}} \psi_2(t, R) q(y(t, R)) dt \geq 0. \tag{14}$$

As we work on compact sets, the test functions do not have to vanish at infinity. However new terms $y(t, R)$ and $y(t, L)$ now appear. Related to this notion of solutions on compact sets, we also have a similar definition for mv entropy solution.

Definition 7 (Measure-valued entropy solution on compact sets). *A Young measure $\mu : (t, x) \in \mathbf{T} \times \mathbf{X} \mapsto \mu_{(t,x)} \in \mathscr{P}(\mathbf{Y})$ is an entropy measure-valued solution to (12) if, for all test functions $\psi_1 \in \mathscr{C}^1(\mathbf{T} \times \mathbf{X})$, it satisfies*

$$\int_{\mathbf{T}} \int_{\mathbf{X}} \left(\frac{\partial \psi_1}{\partial t} \langle \mu_{(t,x)}, y \rangle + \frac{\partial \psi_1}{\partial x} \langle \mu_{(t,x)}, f(y) \rangle \right) dx \, dt + \int_{\mathbf{X}} \psi_1(0, x) \langle \sigma_0, y \rangle dx$$
$$- \int_{\mathbf{X}} \psi_1(T, x) \langle \sigma_T, y \rangle dx + \int_{\mathbf{T}} \psi_1(t, L) \langle \sigma_L, f(y) \rangle dt - \int_{\mathbf{T}} \psi_1(t, R) \langle \sigma_R, f(y) \rangle dt = 0 \tag{15}$$

7

and, for all convex pairs (η, q) and all non-negative test functions $\psi_2 \in \mathscr{C}^1(\mathbf{T} \times \mathbf{X})$, it satisfies

$$\int_{\mathbf{T}} \int_{\mathbf{X}} \left(\frac{\partial \psi_2}{\partial t} \langle \mu_{(t,x)}, \eta(y) \rangle + \frac{\partial \psi_2}{\partial x} \langle \mu_{(t,x)}, q(y) \rangle \right) dx\, dt + \int_{\mathbf{X}} \psi_2(0, x) \langle \sigma_0, \eta(y) \rangle dx$$

$$- \int_{\mathbf{X}} \psi_2(T, x) \langle \sigma_T, \eta(y) \rangle dx + \int_{\mathbf{T}} \psi_2(t, L) \langle \sigma_L, q(y) \rangle dt - \int_{\mathbf{T}} \psi_2(t, R) \langle \sigma_R, q(y) \rangle dt \geq 0,$$

$$\tag{16}$$

where σ_0, σ_T, σ_L resp. σ_R are Young measures supported on $\mathbf{T}$, $\mathbf{X}$, $\mathbf{T}$ resp. $\mathbf{X}$.

Remark 3 (Imposing constraints on the boundary)**.** *To ensure the concentration of $\mu_{(t,x)}$ on the graph of the solution to (13)-(14), in addition to the condition $\sigma_0 = \delta_{y_0(x)}$, one may impose conditions on the boundary measures σ_L and/or σ_R. In practice, one knows the initial condition in an interval larger than $\mathbf{X}$ and so one is able to impose σ_L and/or σ_R. The width of this interval depends on the Lipschitz constant of the flux, T, L and R. As an illustrative example, consider the case where the initial condition is positive and the flux is strictly convex. By the classical method of characteristics, if the initial condition y_0 is positive then so is the solution y for all $t \geq 0$. In particular if f is strictly convex we only need to impose knowledge at the left of the box $\mathbf{X}$. Therefore σ_L has to be known for all $t \in \mathbf{T}$, and σ_R is unconstrained. We refer to [27] for a more precise discussion on the choice of the boundary constraint.*

3 A convex optimization approach for mv solutions on compact sets

In the latter section, we introduced mv solutions for scalar hyperbolic equations. Note that measures are fully characterized by their moments on compact sets, see e.g. [24, p. 52]. This means in particular that moments are the quantities of interest. The aim of this section is to express formulations (15)-(16) as constraints on the moments, to explain how one can compute numerically these moments thanks to the so-called moment-SOS hierarchy. We also show how one can interpret these moments in the case where the initial measure in concentrated.

3.1 Moment constraints for the entropy mv solution

Let $\nu \in \mathscr{M}(\mathbf{K})_+$, with $\mathbf{K} := \mathbf{T} \times \mathbf{X} \times \mathbf{Y}$. In the following, we derive moment constraints that will imply that ν can be desintegrated as follows

$$d\nu(t, x, y) = dt\, dx\, d\mu_{(t,x)}(dy) \tag{17}$$

or, equivalently,

$$\nu = \lambda_{\mathbf{T}} \lambda_{\mathbf{X}} \mu_{(t,x)}, \tag{18}$$

where μ is an entropy mv solution satisfying (15) and (16). In (17) the measure ν is called an *occupation measure* and the Young measure μ is its conditional measuring y given t and x. We also need to introduce the following time boundary measures

$$d\nu_0(t, x, y) := \delta_0(dt)\, dx\, \sigma_0(dy), \quad d\nu_T(t, x, y) := \delta_T(dt)\, dx\, \sigma_T(dy) \tag{19}$$

8

whose supports are $\mathbf{K}_0 := \{0\} \times \mathbf{X} \times \mathbf{Y}$ and $\mathbf{K}_T := \{T\} \times \mathbf{X} \times \mathbf{Y}$ respectively. Similarly, we introduce the following space boundary measures.

$$d\nu_L(t, x, y) := dt \, \delta_L(dx) \, \sigma_L(dy), \quad d\nu_R(t, x, y) := dt \, \delta_R(dx) \, \sigma_R(dy) \tag{20}$$

whose supports are given by $\mathbf{K}_L := \mathbf{T} \times \{L\} \times \mathbf{Y}$ and $\mathbf{K}_R := \mathbf{T} \times \{R\} \times \mathbf{Y}$ respectively.

First, to ensure that the marginal of ν with respect to t and x is the Lebesgue measure on $\mathbf{T} \times \mathbf{X}$, it suffices to impose that:

$$\int_{\mathbf{K}} t^{\alpha_1} x^{\alpha_2} \, d\nu(t, x, y) = \int_{\mathbf{T} \times \mathbf{X}} t^{\alpha_1} x^{\alpha_2} \, dt \, dx, \quad \alpha \in \mathbb{N}^2. \tag{21}$$

In a similar manner, we can enforce the respective marginal of the boundary measures to be Lebesgue as follows

$$\int_{\mathbf{K}_0} 0^{\alpha_1} x^{\alpha_2} \, d\nu_0(t, x, y) = \int_{\mathbf{X}} 0^{\alpha_1} x^{\alpha_2} dx, \quad \alpha \in \mathbb{N}^2, \tag{22}$$

$$\int_{\mathbf{K}_T} T^{\alpha_1} x^{\alpha_2} \, d\nu_T(t, x, y) = \int_{\mathbf{X}} T^{\alpha_1} x^{\alpha_2} dx, \quad \alpha \in \mathbb{N}^2, \tag{23}$$

$$\int_{\mathbf{K}_L} t^{\alpha_1} L^{\alpha_2} \, d\nu_L(t, x, y) = \int_{\mathbf{T}} t^{\alpha_1} L^{\alpha_2} dt, \quad \alpha \in \mathbb{N}^2 \tag{24}$$

and

$$\int_{\mathbf{K}_R} t^{\alpha_1} R^{\alpha_2} \, d\nu_R(t, x, y) = \int_{\mathbf{T}} t^{\alpha_1} R^{\alpha_2} dt, \quad \alpha \in \mathbb{N}^2. \tag{25}$$

Next, we aim at proving that (15) and (16) can also be expressed by moment constraints. We split the exposition into two steps: the first one deals with (15), while the second one deals with (16).

3.1.1 First step: enforcing (15) by moment constraints

Lemma 1. *Let $\phi_1^\alpha(t, x, y) := t^{\alpha_1} x^{\alpha_2} y$ and $\phi_2^\alpha(t, x, y) := t^{\alpha_1} x^{\alpha_2} f(y)$ for $\alpha \in \mathbb{N}^2$. Linear constraint (15) is equivalent to*

$$\int_{\mathbf{K}} \left(\frac{\partial \phi_1^\alpha}{\partial t} y + \frac{\partial \phi_2^\alpha}{\partial x} f(y) \right) d\nu + \int_{\mathbf{K}_0} \phi_1^\alpha \, d\nu_0 - \int_{\mathbf{K}_T} \phi_1^\alpha \, d\nu_T \\ + \int_{\mathbf{K}_L} \phi_2^\alpha \, d\nu_L - \int_{\mathbf{K}_R} \phi_2^\alpha \, d\nu_R = 0 \tag{26}$$

for all $\alpha \in \mathbb{N}^2$.

Proof of Lemma 1 Since $\mathbf{T} \times \mathbf{X}$ is a compact set, as a consequence of the Stone-Weierstrass theorem, we can restrict the test functions to $\psi_1 = t^{\alpha_1} x^{\alpha_2}$ for $\alpha \in \mathbb{N}^2$ to enforce (15). $\square$

3.1.2 Second step: enforcing (16) by moment constraints

As noticed in Remark 2, the entropy inequality is satisfied for all the convex pairs (η, q) if and only it is satisfied for all Kruzhkov entropies given in (8). To express (16) as moment constraints, we are faced with two issues: first, taking into account an uncountable family of functions parametrized by $v \in \mathbf{Y}$ and, second, the absolute value function $v \mapsto |v|$ is not a polynomial. To deal with the uncountable family of functions, we introduce v as a new variable. To treat the absolute value, we double the number of measures.

More precisely, we define the Borel measures ϑ^+, ϑ^- whose supports are defined as follows

$$\mathbf{K}^+ := \mathrm{spt}(\vartheta^+) = \{(t, x, y, v) \in \mathbf{K} \times \mathbf{Y} : y \geq v\},$$

$$\mathbf{K}^- := \mathrm{spt}(\vartheta^-) = \{(t, x, y, v) \in \mathbf{K} \times \mathbf{Y} : y \leq v\}.$$

Similarly, we define the time boundary measures ϑ_0^+, ϑ_0^-, ϑ_T^+ and ϑ_T^- with the following supports

$$\begin{aligned}
\mathbf{K}_0^+ &:= \mathrm{spt}(\vartheta_0^+) = \{(t, x, y, v) \in \mathbf{K}_0 \times \mathbf{Y} : y \geq v\}, \\
\mathbf{K}_0^- &:= \mathrm{spt}(\vartheta_0^-) = \{(t, x, y, v) \in \mathbf{K}_0 \times \mathbf{Y} : y \leq v\},
\end{aligned} \tag{27}$$

and

$$\begin{aligned}
\mathbf{K}_T^+ &:= \mathrm{spt}(\vartheta_T^+) = \{(t, x, y, v) \in \mathbf{K}_T \times \mathbf{Y} : y \geq v\}, \\
\mathbf{K}_T^- &:= \mathrm{spt}(\vartheta_T^-) = \{(t, x, y, v) \in \mathbf{K}_T \times \mathbf{Y} : y \leq v\}.
\end{aligned} \tag{28}$$

Finally, let us define the space boundary measures ϑ_L^+, ϑ_L^-, ϑ_R^+ and ϑ_R^- with the following supports

$$\begin{aligned}
\mathbf{K}_L^+ &:= \mathrm{spt}(\vartheta_L^+) = \{(t, x, y, v) \in \mathbf{K}_L \times \mathbf{Y} : y \geq v\} \\
\mathbf{K}_L^- &:= \mathrm{spt}(\vartheta_L^-) = \{(t, x, y, v) \in \mathbf{K}_L \times \mathbf{Y} : y \leq v\}
\end{aligned} \tag{29}$$

and

$$\begin{aligned}
\mathbf{K}_R^+ &:= \mathrm{spt}(\vartheta_R^+) = \{(t, x, y, v) \in \mathbf{K}_R \times \mathbf{Y} : y \geq v\}, \\
\mathbf{K}_R^- &:= \mathrm{spt}(\vartheta_R^-) = \{(t, x, y, v) \in \mathbf{K}_R \times \mathbf{Y} : y \leq v\}.
\end{aligned} \tag{30}$$

We are now in position to state the following lemma.

Lemma 2 (Recovering all Kruzhkov entropies)**.** *Assume that*

$$\vartheta^+ + \vartheta^- = \nu \otimes \lambda_{\mathbf{Y}}, \tag{31}$$

$$\vartheta_0^+ + \vartheta_T^- = \nu_0 \otimes \lambda_{\mathbf{Y}}, \quad \vartheta_T^+ + \vartheta_T^- = \nu_T \otimes \lambda_{\mathbf{Y}}, \tag{32}$$

$$\vartheta_L^+ + \vartheta_L^- = \nu_L \otimes \lambda_{\mathbf{Y}}, \quad \vartheta_R^+ + \vartheta_R^- = \nu_R \otimes \lambda_{\mathbf{Y}}. \tag{33}$$

Then, (16) *is equivalent to*

$$\int_{\mathbf{K}^+} \theta(v) \left(\frac{\partial \psi_2}{\partial t}(y - v) + \frac{\partial \psi_2}{\partial x}(f(y) - f(v)) \right) d\vartheta^+$$

$$+ \int_{\mathbf{K}^-} \theta(v) \left(\frac{\partial \psi_2}{\partial t}(v - y) + \frac{\partial \psi_2}{\partial x}(f(v) - f(y)) \right) d\vartheta^-$$

$$+ \int_{\mathbf{K}_0^+} \theta(v)\psi_2(0, x)(y - v)d\vartheta_0^+ + \int_{\mathbf{K}_0^-} \theta(v)\psi_2(0, x)(v - y)d\vartheta_0^-$$

$$- \int_{\mathbf{K}_T^+} \theta(v)\psi_2(T, x)(y - v)d\vartheta_T^+ - \int_{\mathbf{K}_T^-} \theta(v)\psi_2(T, x)(v - y)d\vartheta_T^- \tag{34}$$

$$+ \int_{\mathbf{K}_L^+} \theta(v)\psi_2(t, L)(f(y) - f(v))d\vartheta_L^+ + \int_{\mathbf{K}_L^-} \theta(v)\psi_2(t, L)(f(v) - f(y))d\vartheta_L^-$$

$$- \int_{\mathbf{K}_R^+} \theta(v)\psi_2(t, R)(f(y) - f(v))d\vartheta_R^+ - \int_{\mathbf{K}_R^-} \theta(v)\psi_2(t, R)(f(v) - f(y))d\vartheta_R^- \geq 0,$$

for all nonnegative functions $\theta \in \mathscr{C}(\mathbf{Y})$.

Note that from the Stone-Weierstrass Theorem, the constraints (31), (32) and (33) can be expressed as moment constraints: (31) holds if and only if, for all $\alpha \in \mathbb{N}^4$,

$$\int_{\mathbf{K}^+ \cup \mathbf{K}^-} t^{\alpha_1} x^{\alpha_2} y^{\alpha_3} v^{\alpha_4} d(\vartheta^+ + \vartheta^-)(t, x, y, v) = \int_{\mathbf{K}} t^{\alpha_1} x^{\alpha_2} y^{\alpha_3} d\nu(t, x, y) \int_{\mathbf{Y}} v^{\alpha_4} dv \tag{35}$$

and similarly for (32) and (33).

Proof of Lemma 2: For conciseness, we deal only with the constraint (31). The other constraints of Lemma 2 follow similarly. Let us prove that if for all nonnegative functions $\theta \in \mathscr{C}^1(\mathbf{Y})$ and all nonnegative functions $\psi_2 \in \mathscr{C}^1(\mathbf{T} \times \mathbf{X})$,

$$\int_{\mathbf{K}^+} \theta(v) \left(\frac{\partial \psi_2}{\partial t}(y - v) + \frac{\partial \psi_2}{\partial x}(f(y) - f(v)) \right) d\vartheta^+$$

$$+ \int_{\mathbf{K}^-} \theta(v) \left(\frac{\partial \psi_2}{\partial t}(v - y) + \frac{\partial \psi_2}{\partial x}(f(v) - f(y)) \right) d\vartheta^- = \tag{36}$$

$$\int_{\mathbf{K}^+ \cup \mathbf{K}^-} \theta(v) \left(\frac{\partial \psi_2}{\partial t}|y - v| + \frac{\partial \psi_2}{\partial x}\text{sign}(y - v)(f(y) - f(v)) \right) d(\vartheta^+ + \vartheta^-) \geq 0$$

then the following inequality holds, for all test functions $\psi_2 \in \mathscr{C}^1(\mathbf{T} \times \mathbf{X})$ and all $v \in \mathbf{Y}$:

$$\int_{\mathbf{T} \times \mathbf{X} \times \mathbf{Y}} \frac{\partial \psi_2}{\partial t}|y - v| + \frac{\partial \psi_2}{\partial x}\text{sign}(y - v)(f(y) - f(v))d\nu \geq 0. \tag{37}$$

First, observe that (31) implies that

$$\int_{\mathbf{K}^+ \cup \mathbf{K}^-} \theta(v) \left(\frac{\partial \psi_2}{\partial t}|y - v| + \frac{\partial \psi_2}{\partial x}\text{sign}(y - v)(f(y) - f(v)) \right) d(\vartheta^+ + \vartheta^-) =$$

$$\int_{\mathbf{Y}} \theta(v) \left(\int_{\mathbf{T} \times \mathbf{X} \times \mathbf{Y}} \left(\frac{\partial \psi_2}{\partial t}|y - v| + \frac{\partial \psi_2}{\partial x}\text{sign}(y - v)(f(y) - f(v)) \right) d\nu \right) dv. \tag{38}$$

Then, since (36) holds for any nonnegative functions θ, and $y - v = |y - v|$ on $\text{spt}(\vartheta^+)$ (resp. $v - y = |y - v|$ on $\text{spt}(\vartheta^-)$),

$$\int_{\mathbf{T} \times \mathbf{X} \times \mathbf{Y}} \left(\frac{\partial \psi_2}{\partial t}|y - v| + \frac{\partial \psi_2}{\partial x}\text{sign}(y - v)(f(y) - f(v)) \right) d\nu \geq 0. \tag{39}$$

11

$\square$

In order to express (16) as moment constraints, it remains to prove that the functions ψ_2 and θ can be replaced by suitable polynomials. Here, in contrast with the first step, where the functions ψ_1 were unconstrained, the functions ψ_2 and θ have to be nonnegative. To address this issue, we again use positivity certificates from real algebraic geometry.

Lemma 3. *Let*

$$
\begin{aligned}
\phi_1^{+,\alpha}(t,x,y,v) &:= t^{\alpha_1}(T-t)^{\alpha_2}(x-L)^{\alpha_3}(R-x)^{\alpha_4}(v-\underline{y})^{\alpha_5}(\bar{y}-v)^{\alpha_6}(y-v), \\
\phi_1^{-,\alpha}(t,x,y,v) &:= t^{\alpha_1}(T-t)^{\alpha_2}(x-L)^{\alpha_3}(R-x)^{\alpha_4}(v-\underline{y})^{\alpha_5}(\bar{y}-v)^{\alpha_6}(v-y), \\
\phi_2^{+,\alpha}(t,x,y,v) &::= t^{\alpha_1}(T-t)^{\alpha_2}(x-L)^{\alpha_3}(R-x)^{\alpha_4}(v-\underline{y})^{\alpha_5}(\bar{y}-v)^{\alpha_6}(f(y)-f(v)), \\
\phi_2^{-,\alpha}(t,x,y,v) &:= t^{\alpha_1}(T-t)^{\alpha_2}(x-L)^{\alpha_3}(R-x)^{\alpha_4}(v-\underline{y})^{\alpha_5}(\bar{y}-v)^{\alpha_6}(f(v)-f(y))
\end{aligned}
\tag{40}
$$

for $\alpha \in \mathbb{N}^6$. Then, (16) is equivalent to

$$
\begin{aligned}
&\int_{\mathbf{K}^+}\left(\frac{\partial \phi_1^{+,\alpha}}{\partial t}+\frac{\partial \phi_2^{+,\alpha}}{\partial x}\right)d\vartheta^+ + \int_{\mathbf{K}^-}\left(\frac{\partial \phi_1^{-,\alpha}}{\partial t}+\frac{\partial \phi_2^{-,\alpha}}{\partial x}\right)d\vartheta^- \\
&+\int_{\mathbf{K}_0^+}\phi_1^{+,\alpha}d\vartheta_0^+ + \int_{\mathbf{K}_0^-}\phi_1^{-,\alpha}\,d\vartheta_0^- - \int_{\mathbf{K}_T^+}\phi_1^{+,\alpha}d\vartheta_T^+ - \int_{\mathbf{K}_T^-}\phi_1^{-,\alpha}\,d\vartheta_T^- \\
&+\int_{\mathbf{K}_L^+}\phi_2^{+,\alpha}d\vartheta_L^+ + \int_{\mathbf{K}_L^-}\phi_2^{-,\alpha}d\vartheta_L^- - \int_{\mathbf{K}_R^+}\phi_2^{+,\alpha}d\vartheta_R^+ - \int_{\mathbf{K}_R^-}\phi_2^{-,\alpha}d\vartheta_R^- \geq 0
\end{aligned}
\tag{41}
$$

for all $\alpha \in \mathbb{N}^6$.

Proof of Lemma 3: The proof relies on a result of real algebraic geometry. Again, invoking the Stone-Weierstrass Theorem, in (16) we can restrict the test functions ψ_2 and θ to be polynomials. To enforce their positivity, we use Handelman's Positivstellensatz [19] that implies that

$$
\begin{aligned}
\psi_2(t,x) &= \sum_{\alpha_1,\alpha_2,\alpha_3,\alpha_4\in\mathbb{N}^4} c_\alpha^{\psi_2} t^{\alpha_1}(T-t)^{\alpha_2}(x-L)^{\alpha_3}(R-x)^{\alpha_4}, \\
\theta(v) &= \sum_{\alpha_5,\alpha_5\in\mathbb{N}} c_\alpha^{\theta}(v-\underline{y})^{\alpha_5}(\bar{y}-v)^{\alpha_6}
\end{aligned}
\tag{42}
$$

with finitely many positive real coefficients $c_\alpha^{\psi_2}, c_\alpha^{\theta}$. Now, (41) implies that

$$
\begin{aligned}
\sum_{\alpha\in\mathbb{N}^6} c_\alpha^{\psi_2} c_\alpha^{\theta} &\left\{\int_{\mathbf{K}^+}\left(\frac{\partial \phi_1^{+,\alpha}}{\partial t}+\frac{\partial \phi_2^{+,\alpha}}{\partial x}\right)d\vartheta^+ + \int_{\mathbf{K}^-}\left(\frac{\partial \phi_1^{-,\alpha}}{\partial t}+\frac{\partial \phi_2^{-,\alpha}}{\partial x}\right)d\vartheta^-\right. \\
&+\int_{\mathbf{K}_0^+}\phi_1^{+,\alpha}d\vartheta_0^+ + \int_{\mathbf{K}_0^-}\phi_1^{-,\alpha}d\vartheta_0^- - \int_{\mathbf{K}_T^+}\phi_1^{+,\alpha}d\vartheta_T^+ - \int_{\mathbf{K}_T^-}\phi_1^{-,\alpha}d\vartheta_T^- \\
&\left.+\int_{\mathbf{K}_L^+}\phi_2^{+,\alpha}d\vartheta_L^+ + \int_{\mathbf{K}_L^-}\phi_2^{-,\alpha}d\vartheta_L^- - \int_{\mathbf{K}_R^+}\phi_2^{+,\alpha}d\vartheta_R^+ - \int_{\mathbf{K}_R^-}\phi_2^{-,\alpha}d\vartheta_R^-\right\} \geq 0,
\end{aligned}
\tag{43}
$$

12

which, by linearity of the integrals and the derivatives, recovers (34) for ψ_2 and θ given in (42). Consequently, by Lemma 2, (34) implies that (16) holds. $\qquad\square$

Remark 4. *Note that the measure defined in* (17) *is similar to the occupation measure introduced in [25], which deals with optimal control of nonlinear finite-dimensional systems. This notion has been further used in many other contexts, as for instance the computation of region of attraction [20]. Therefore, the formulation given in* (15)-(16) *might be instrumental to solve other problems than computing numerically the solution of scalar hyperbolic PDE.*

The next section aims at showing that a moment formulation can be numerically solved thanks to the moment-SOS hierarchy and SDP.

3.2 The Generalized Moment Problem and its relaxations

Roughly speaking, the *Generalized Moment Problem* (GMP) is an infinite-dimensional linear optimization problem on finitely many Borel measures $\nu_i \in \mathscr{M}(\mathbf{K}_i)_+$ whose supports are contained in given sets $\mathbf{K}_i \subseteq \mathbb{R}^{n_i}$, with $i = 1, \ldots, k$ and $n_i \in \mathbb{N}$. That is, one is interested in finding measures whose moments satisfy (possibly countably many) linear constraints and which minimize a linear criterion. In full generality, the GMP is intractable, but if all $\mathbf{K}_i$ are basic semi-algebraic sets and the integrands are polynomials (semi-algebraic functions are also allowed[1]), then one may provide an efficient numerical scheme to approximate as closely as desired any finite number of moments of optimal solutions of the GMP. It consists of solving a hierarchy of semidefinite programs[2] of increasing size. Convergence of this numerical scheme is guaranteed by invoking powerful results from Real Algebraic Geometry (essentially positivity certificates).

Let $h_i \in \mathbb{R}[\mathbf{w}^i]$ and $h_{i,k} \in \mathbb{R}[\mathbf{w}^i]$ be polynomials in the vector of indeterminates $\mathbf{w}^i \in \mathbb{R}^{n_i}$ and let b_k be real numbers, for finitely many $i = 1, 2, \ldots, N$ and countably many $k = 1, 2, \ldots$. The GMP can be formulated as

$$
\begin{aligned}
\rho^\star := \quad \inf_\nu \quad & \sum_{i=1}^N \int_{\mathbf{K}_i} h_i d\nu_i \\
\text{s.t.} \quad & \sum_{i=1}^N \int_{\mathbf{K}_i} h_{i,k} d\nu_i \leqq b_k, \quad k = 1, 2, \ldots \\
& \nu_i \in \mathscr{M}(\mathbf{K}_i)_+, \quad i = 1, \ldots, N.
\end{aligned}
\tag{44}
$$

Entropy mv solution as a GMP In the scalar hyperbolic case, the measures ν_i under consideration are $\nu, \nu_T, \nu_0, \nu_R, \nu_L$ and all the measures we have introduced when transforming the Kruzhkov inequality into moment constraints. The sets $\mathbf{K}_i$ correspond to $\mathbf{T}$, $\mathbf{X}$ and $\mathbf{Y}$. Finally, the polynomials $h_{i,j}$ are given in (26) (conservation law) (41) (entropy inequality), (35) (Kruzhkov entropies), and (21-25) (boundary measures).

We may also define an objective functional

$$
\int_{\mathbf{K}} h d\nu + \int_{\mathbf{K}_0} h_0 d\nu_0 + \int_{\mathbf{K}_T} h_T d\nu_T + \int_{\mathbf{K}_L} h_L d\nu_L + \int_{\mathbf{K}_R} h_R d\nu_R,
\tag{45}
$$

---

[1]A semi-algebraic function is a function whose graph is a semi-algebraic set, i.e. it is described by finitely many polynomial inequalities and equations.

[2]A semidefinite program is a particular class of a convex conic optimization problem that can be solved numerically efficiently.

with $h, h_0, h_T, h_L, h_R \in \mathbb{R}[t, x, y]$.

If $\sigma_0 = \delta_{y_0(x)}$ with y_0 an initial condition in (13)-(14) and, in addition, if one imposes suitable boundary measures as exposed in Remark 3, then this objective functional is not especially useful to recover the entropy mv solution of scalar hyperbolic PDE, since the corresponding Young measure is concentrated as a consequence of Theorem 1: there is nothing to be optimized. However, with such an objective functional, one can compute quantities of interest such as the energy of the solution. Moreover, our aim is to *relax* the GMP in order to solve it numerically and, then, this objective functional might be helpful to accelerate the convergence of the corresponding relaxations. We refer to Section 4 for more discussions about the choice of objective functionals for the Riemann problem of the Burgers equation.

Finally, one is able to define a GMP describing entropy mv solution:

$$
\begin{aligned}
&\inf_{\nu, \, \nu_T, \, \nu_L \text{ and/or } \nu_R} && \text{(45) (objective functional)} \\
&\text{s.t.} && \text{(26) (conservation law),} \\
& && \text{(41) (entropy inequality),} \\
& && \text{(35) (Kruzhkov entropies),} \\
& && (21-25) \text{ (boundary measures),} \\
& && \nu \in \mathcal{M}(\mathbf{K})_+ \text{ (occupation measure),} \\
& && \nu_T \in \mathcal{M}(\mathbf{K}_T)_+ \text{ (time boundary measure),} \\
& && \nu_L \in \mathcal{M}(\mathbf{K}_L)_+ \text{ (space boundary measure),} \\
& && \text{and/or } \nu_R \in \mathcal{M}(\mathbf{K}_R)_+ \text{ (space boundary measures)}
\end{aligned}
\qquad (46)
$$

where the measures defined in (31)-(33) and related to the Kruzhkov entropies are considered as implicit variables.

From measures to moments   Instead of optimizing over the measures in problem (46), we optimize over their moments. For simplicity and clarity of exposition, we describe the approach in the case of a single unknown measure ν, but it easily extends to the case of several measures. So consider the simplified GMP:

$$
\begin{aligned}
\rho^\star = \quad &\inf_\nu && \int_{\mathbf{K}} h \, d\nu \\
&\text{s.t.} && \int_{\mathbf{K}} h_k \, d\nu \leqq b_k, \quad k = 1, 2, \dots \\
& && \nu \in \mathcal{M}(\mathbf{K})_+.
\end{aligned}
\qquad (47)
$$

The *moment sequence* $\mathbf{z} = (\mathbf{z}_\alpha)_{\alpha \in \mathbb{N}^n}$ of a given measure $\nu \in \mathcal{M}(\mathbf{K})_+$ is defined by

$$
\mathbf{z}_\alpha = \int_K \mathbf{w}^\alpha \, d\nu, \quad \alpha \in \mathbb{N}^n
\qquad (48)
$$

where $\mathbf{w}^\alpha = w_1^{\alpha_1} \cdot \ldots \cdot w_n^{\alpha_n}$. Conversely, given a sequence $(\mathbf{z}_\alpha)_{\alpha \in \mathbb{N}^n}$, if (48) holds for some $\nu \in \mathcal{M}(\mathbf{K})_+$ we say that the sequence has the representing measure. Recall that measures on compact sets are uniquely characterized by their moments; see e.g. [24, p. 52].

Let $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \leq d\}$, where $|\alpha| := \sum_{i=1}^n \alpha_i$, and $s(d) := \binom{n+d}{d}$. A vector $\mathbf{p} := (\mathbf{p}_\alpha)_{\alpha \in \mathbb{N}_d^n} \in \mathbb{R}^{s(d)}$ with $d = \deg(p)$ is the coefficient vector (in the monomial basis) of a polynomial $p \in \mathbb{R}[\mathbf{w}]$ expressed as $p = \sum_{\alpha \in \mathbb{N}_{\deg(p)}^n} \mathbf{p}_\alpha \mathbf{w}^\alpha$. Next, integration of p with respect to a measure ν involves only finitely many moments:

$$
\int_{\mathbf{K}} p \, d\nu = \int_{\mathbf{K}} \sum_{\alpha \in \mathbb{N}_d^n} \mathbf{p}_\alpha \mathbf{w}^\alpha \, d\nu = \sum_{\alpha \in \mathbb{N}_d^n} \mathbf{p}_\alpha \int_{\mathbf{K}} \mathbf{w}^\alpha \, d\nu = \sum_{\alpha \in \mathbb{N}_d^n} \mathbf{p}_\alpha \mathbf{z}_\alpha.
$$

Next, define a pseudo-integration with respect to an arbitrary sequence $\mathbf{z} \in \mathbb{R}^{\mathbb{N}^n}$ by:

$$\ell_{\mathbf{z}}(p) := \sum_{\alpha \in \mathbb{N}^n} \mathbf{p}_\alpha z_\alpha \tag{49}$$

and $\ell_{\mathbf{z}}$ is called the *Riesz functional*. Moment sequences can be characterized via the Riesz functional:

Theorem 2 (Riesz-Haviland [24, Theorem 3.1]). *Let $\mathbf{K} \subseteq \mathbb{R}^n$ be closed. A real sequence $\mathbf{z}$ is the sequence of some measure ν supported on $\mathbf{K}$ if and only if $\ell_{\mathbf{z}}(p) \geq 0$ for all $p \in \mathbb{R}[\mathbf{w}]$ nonnegative on $\mathbf{K}$.*

Assuming that $\mathbf{K}$ is closed, we can reformulate the GMP (47) as a linear problem on moment sequences. Consider the optimization problem:

$$\begin{aligned}
\rho^* = \quad &\inf_{\mathbf{z}} \quad \ell_{\mathbf{z}}(h) \\
&\text{s.t.} \quad \ell_{\mathbf{z}}(h_k) \lesseqgtr b_k, \ k = 1, 2, \dots \\
&\qquad \ell_{\mathbf{z}}(p) \geq 0, \text{ for all } p \in \mathbb{R}[\mathbf{w}] \text{ positive on } \mathbf{K}.
\end{aligned} \tag{50}$$

By Theorem 2, the two formulations (50) and (47) are equivalent. Of course problem (50) is still numerically intractable.

The second and last step to approximate GMPs numerically consists of replacing the cone of polynomials positive on $\mathbf{K}$ by a more tractable cone. This is where one exploits the fact that $\mathbf{K}$ is basic semi-algebraic set.

From nonnegative polynomials to sums of squares Characterizing nonnegativity of polynomials is an important issue in real algebraic geometry. Let $\mathbf{K}$ be a *basic semi-algebraic set*, i.e.:

$$\mathbf{K} = \{\mathbf{w} \in \mathbb{R}^n : g_1(\mathbf{w}) \geq 0, \dots, g_m(\mathbf{w}) \geq 0\}. \tag{51}$$

for some polynomials $g_1, \dots, g_m \in \mathbb{R}[\mathbf{w}]$, and assume that $\mathbf{K}$ is compact. In addition assume that one of the polynomials, i.e. the first one, is $g_1(\mathbf{w}) := N - \sum_{i=1}^n \mathbf{w}_i^2$ for some $N \in \mathbb{N}$ sufficiently large[3]. For notational convenience we let $g_0(\mathbf{w}) := 1$.

Remark 5. *Note that the compact sets $\mathbf{T}, \mathbf{X}$ and $\mathbf{Y}$ which are defined in the latter section can be expressed as basic semi-algebraic compact sets. Indeed, one has*

$$\begin{aligned}
&\mathbf{T} = \{t \in \mathbb{R} : t(T-t) \geq 0\}, \ \mathbf{X} = \{x \in \mathbb{R} : (x-L)(R-x) \geq 0\}, \\
&\mathbf{Y} = \{y \in \mathbb{R} : (y-\underline{y})(\bar{y}-y) \geq 0\}.
\end{aligned} \tag{52}$$

Recall that a polynomial $s \in \mathbb{R}[\mathbf{w}]$ is a *sum of squares* (SOS) if there are finitely many polynomials $q_1, \dots, q_r$ such that $s(\mathbf{w}) = \sum_{j=1}^r q_j(\mathbf{w})^2$ for all $\mathbf{w}$.

The following result due to Putinar [32] is crucial to approximate (50) numerically.

Theorem 3 (Putinar's Positivstellensatz). *If $p > 0$ on $\mathbf{K}$ then $p = \sum_{j=0}^m s_j g_j$ for some SOS polynomials $s_j \in \mathbb{R}[\mathbf{w}]$, $j = 0, 1, \dots, m$.*

---

[3]This condition is slightly stronger than asking $\mathbf{K}$ to be basic semi-algebraic compact. However, the inequality $N - \sum_{i=1}^n \mathbf{w}_i^2 \geq 0$ can always be added as a redundant constraint to the description of a basic semi-algebraic compact set.

By a density argument, checking nonnegativity of $\ell_{\mathbf{z}}$ on polynomials positive on $\mathbf{K}$ can be replaced by checking nonnegativity only on polynomials that are strictly positive on $\mathbf{K}$ and hence on those that have an SOS representation as in Theorem 3.

Next, for a given integer d, denote by $\Sigma[\mathbf{w}]_d \subset \mathbb{R}[\mathbf{w}]$ the set of SOS polynomials of degree at most $2d$, and define the cone $Q(g) \subset \mathbb{R}[\mathbf{w}]$ by:

$$Q_d(g) := \left\{ \sum_{j=0}^{m} \sigma_j \, g_j : \, \deg(\sigma_j \, g_j) \leq 2d, \, \sigma_j \in \Sigma[\mathbf{w}], \, j = 0, 1, \ldots, m \right\} \qquad (53)$$

and observe that $Q_d(g) \subset Q_{d+1}(g)$ consist of polynomials positive on $\mathbf{K}$ for all d.

Let $\mathbf{v}_d := (\mathbf{w}^\alpha)_{|\alpha| \leq d} \in \mathbb{R}[\mathbf{w}]^{s(d)}$ be the vector of monomials of degree at most d. For instance, for $n = 2$ and $d = 3$, $\mathbf{v}_d = (1 \; w_1 \; w_2 \; w_1^2 \; w_1 w_2 \; w_2^2 \; w_1^3 \; w_1^2 w_2 \; w_1 w_2^2 \; w_2^3)$. For $j = 0, 1, \ldots, m$, let d_j denote the smallest integer larger than or equal to $\deg(g_j)$, let $M_{d-d_j}(g_j \, \mathbf{z})$ denote the real symmetric matrix linear in $\mathbf{z}$ corresponding to the entrywise application of $\ell_{\mathbf{z}}$ to the matrix polynomial $g_j \, \mathbf{v}_{d-d_j} \mathbf{v}_{d-d_j}^T$. It turns out that $\ell_{\mathbf{z}}(g_j \, q^2) \geq 0$ for all $q \in \mathbb{R}[\mathbf{w}]_d$ if and only if $M_{d-d_j}(g_j \, \mathbf{z}) \succeq 0$ where the inequality means positive semidefinite. Therefore checking whether $\ell_{\mathbf{z}}$ is nonnegative on $Q_d(g)$ reduces to checking whether $M_{d-d_j}(g_j \, \mathbf{z}) \succeq 0$ for $j = 0, 1, \ldots, m$, which are convex linear matrix inequalities in $\mathbf{z}$.

Moment-SOS hierarchy   The following finite-dimensional semidefinite programming (SDP) problems are relaxations of the moment problem (50):

$$\begin{aligned}
\rho_d^* = \; & \inf_{\mathbf{z}} \quad \ell_{\mathbf{z}}(h) \\
& \text{s.t.} \quad \ell_{\mathbf{z}}(h_k) \lessgtr b_k, \; \deg(h_k) \leq 2d, \; k = 1, 2, \ldots \\
& \qquad\quad M_{d-d_j}(g_j \, \mathbf{z}) \succeq 0, \; j = 0, 1, \ldots, m
\end{aligned} \qquad (54)$$

and they are parametrized by the relaxation order d.

Theorem 4 (Convergence of the moment-SOS hierarchy [24]). *Assume there is some $M > 0$ and $k \in \mathbb{N}$ such that $\ell_{\mathbf{z}}(h_k) \leq b_k$ implies $\ell_{\mathbf{z}}(1) < M$. Then:*

(i) The semidefinite relaxation (54) has an optimal solution $\mathbf{z}^d = (\mathbf{z}_\alpha^d)$, $\rho_d \leq \rho_{d+1}$ and $\lim_{d \to \infty} \rho_d = \rho^*$;

(ii) If (50) has a unique minimizer $\mathbf{z}^*$, then

$$\lim_{d \to \infty} \mathbf{z}_\alpha^d = \mathbf{z}_\alpha^*, \quad \forall \alpha \in \mathbb{N}^n. \qquad (55)$$

A proof is provided in [24], but, for clarity, we recall the steps and the arguments used to obtain the result.

Proof of Theorem 4: Let $\mathbf{z} := (\mathbf{z}_\alpha)_{\alpha \in \mathbb{N}^n_{2d}}$ be a feasible solution of (54). From $M_{d-d_1}(g_1 \, \mathbf{z}) \succeq 0$ we obtain $\sum_{i=1}^{n} \ell_{\mathbf{z}}(\mathbf{w}_i^2) \leq N\mathbf{z}_0$, and in particular $\ell_{\mathbf{z}}(\mathbf{w}_i^2) \leq N\mathbf{z}_0$, $i = 1, \ldots, n$. By iterating one also obtains $\ell(\mathbf{w}_i^{2d}) \leq N^d \mathbf{z}_0$, $i = 1, \ldots, n$. Moreover, combining with $M_d(\mathbf{z}) \succeq 0$, from [24]:

$$|\mathbf{z}_\alpha| \leq \max \left[\mathbf{z}_0, \max_{i=1,\ldots,n} \ell_{\mathbf{z}}(\mathbf{w}_i^{2d}) \right] =: \tau_d, \quad \forall \alpha \in \mathbb{N}^n_{2d}.$$

16

This inequality together with the fact that $\ell_{\mathbf{z}}(1) = \mathbf{z}_0$ is bounded implies that the moment sequence $\mathbf{z}$ is uniformly bounded. Then, the feasible set of (54) is closed, bounded, and hence compact. Hence (54) has an optimal solution $\mathbf{z}^d$.

Next, for a given d, let $\mathbf{z}^d$ be an optimal solution to (54) and complete $\mathbf{z}^d$ with zeros to make it an infinite sequence indexed by $\alpha \in \mathbb{N}^n$. Then define:

$$\hat{\mathbf{z}}^d_\alpha := \frac{\mathbf{z}^d_\alpha}{\tau_k}, \quad \forall \alpha \in \mathbb{N}^n; \quad 2k-1 \leq |\alpha| \leq 2k; \; k = 1, \ldots, d. \tag{56}$$

By construction, $|\hat{\mathbf{z}}^d_\alpha| \leq 1$, for all $\alpha \in \mathbb{N}^n$ and therefore $\hat{\mathbf{z}}^d$ becomes an element of the unit ball $\mathbf{B}_1$ of the Banach space ℓ_∞ of bounded sequences, equipped with the sup-norm. Since ℓ_∞ is the topological dual of ℓ_1, by the Banach-Alaoglu theorem [2, Theorem 3.16], $\mathbf{B}_1$ is weak star (sequentially) compact. Hence there exists $\hat{\mathbf{z}}^\star \in \mathbf{B}_1$ and a subsequence $\{d_k\} \subset \mathbb{N}$ such that $\hat{\mathbf{z}}^{d_k} \to \hat{\mathbf{z}}^\star$ for the weak star topology $\sigma(\ell_\infty, \ell_1)$. In particular, for every $\alpha \in \mathbb{N}^n$, $\lim_{k\to\infty} \hat{\mathbf{z}}^{d_k}_\alpha = \hat{\mathbf{z}}^\star_\alpha$. Since τ_k is bounded for all $k = 1, \ldots, d$, using (56) in the other direction, there exists $\mathbf{z}^* = (\mathbf{z}^*_\alpha)_{\alpha \in \mathbb{N}^n}$ such that

$$\lim_{k\to\infty} \mathbf{z}^{d_k}_\alpha = \mathbf{z}^\star_\alpha, \quad \forall \alpha \in \mathbb{N}^n. \tag{57}$$

The pointwise convergence (57) implies $M_d(g_j\, \mathbf{z}^\star) \succeq 0$ for $j = 0, 1, \ldots, m$. Hence by Theorem 3, $\mathbf{z}^*$ has a representing measure ν supported on $\mathbf{K}$. In particular, from (57), $\ell_{\mathbf{z}}(h_k) \leqq b_k$ for all $k = 1, 2, \ldots$ which proves that ν is a feasible solution of (47). In addition,

$$\rho^\star \geq \lim_{k\to+\infty} \rho^\star_{d_k} = \int_{\mathbf{K}} H d\nu \geq \rho^\star, \tag{58}$$

which proves that ν is an optimal solution of (47). Finally, if (47) has a unique minimizer $\nu^\star$ then $\nu = \nu^\star$ and the convergence (57) holds for the whole sequence, which yields (55). This concludes the proof. □

Remark 6 (Extension to several measures). *Theorem 4 extends naturally to the GMP (44) with finitely many measures. A more detailed discussion is provided in [24, Section 4.5.2 p. 88].*

Convergence of the relaxations of (46)   Problem (46) can be approximated by a hierarchy of semidefinite relaxations as mentioned in Remark 6. Moreover observe that the mass of all measures appearing in (46) is bounded, because their marginals with respect to time and/or space are Lebesgue. Indeed, for instance, $\int_{\mathbf{T}\times\mathbf{X}\times\mathbf{Y}} d\nu = \int_{\mathbf{T}\times\mathbf{X}\times\mathbf{Y}} dt\, dx\, \mu_{(t,x)}(dy) = 1$, where we have used the fact that μ is a Young measure.

Then, according to Remark 6 and Theorem 4, optimal solutions of the moment-SOS hierarchy (54) (adapted to the present context) converge to optimal solutions of (46) as d goes to infinity, in the sense of (55). In particular, one may extract the mv solution of (15) and even obtain the entropy solution of (13)-(14), provided that σ_L and/or σ_R and σ_0 are concentrated.

3.3   Interpretation of the moment solutions

An optimal solution $\mathbf{z}^d$ at step d of the moment-SOS hierarchy of relaxations (54) adapted to the GMP (46), consists of finite sequences of approximate moments, one

for each unknown measure of (46). If one is interested in statistical properties of the mv solution such as its mean or its variance, the moments provide the perfect information, at least for sufficiently large d. However, if one is rather interested in properties of the graph of the entropy solution, a post processing step is required.

An inverse problem Recovering the graph of the solution $\{(t, x, y(t, x)) : t \in \mathbf{T}, x \in \mathbf{X}\} \subset \mathbf{T} \times \mathbf{X} \times \mathbf{Y}$ from the moments of the measure $\nu = \lambda_{\mathbf{T}} \lambda_{\mathbf{X}} \delta_{y(t,x)}$ is an inverse problem whose detailed study is out of the scope of this paper, see e.g. [4] in the context of controlled ODEs. However, we briefly outline here one possible strategy with a formal justification. It turns out that it works surprising well in all our examples of the Burgers equation with or without shock.

Let $\mathbf{w} = (t, x, y)$ and $\mathbf{z}_\alpha = \int_{\mathbf{T} \times \mathbf{X} \times \mathbf{Y}} \mathbf{w}^\alpha \, d\nu$ denote the vector of moments of ν. For any polynomial $p \in \mathbb{R}[\mathbf{w}]_d$ with vector of coefficients $\mathbf{p}$ in the monomial basis, it holds

$$\mathbf{p}^\top M_d(\mathbf{z})\mathbf{p} = \int p^2 \, d\nu.$$

Consequently, if $\mathbf{p}$ is in the kernel of $M_d(\mathbf{z})$, we have that

$$\int_{\mathbf{T} \times \mathbf{X} \times \mathbf{Y}} p^2 \, d\nu = 0.$$

In other words, the support of the measure is contained in the zero level set of every polynomial (whose vector of coefficients is) in the kernel of the moment matrix. Now let us show how to approximate this kernel numerically.

Since the moment matrix is positive semidefinite, it has a spectral decomposition

$$M_d(\mathbf{z}) = PEP^\top \tag{59}$$

where P is an orthonormal matrix whose columns are denoted $\mathbf{p}_i$, $i = 1, 2, \ldots$ and satisfy $\mathbf{p}_i^\top \mathbf{p}_i = 1$ and $\mathbf{p}_i^\top \mathbf{p}_j = 0$ if $i \neq j$, and E is a diagonal matrix whose diagonal entries are eigenvalues $e_{i+1} \geq e_i \geq 0$ of the moment matrix. Each column $\mathbf{p}_i$ is the vector of coefficients in the monomial basis of a polynomial $p_i \in \mathbb{R}[\mathbf{w}]$, so that

$$\mathbf{p}_i^\top M_d(\mathbf{z})\mathbf{p}_i = \int p_i^2 \, d\nu = e_i.$$

The following result shows that the measure is concentrated on a sublevel set of an SOS polynomial constructed from the spectral decomposition of the moment matrix.

Lemma 4 (Concentration inequality). *Let $r \in \mathbb{N}$ and $\beta > 0$. Define*

$$\gamma = \frac{\sum_{i=1}^r e_i}{\beta}$$

and

$$p_{\text{sos}} = \sum_{i=1}^r p_i^2. \tag{60}$$

Then

$$\nu(\{\mathbf{w} : p_{\text{sos}}(\mathbf{w}) \leq \gamma\}) \geq 1 - \beta.$$

The proof of Lemma 4 follows readily from the inequality

$$\nu(\{\mathbf{w} : p_{\text{sos}}(\mathbf{w}) > \gamma\}) \leq \frac{\int p_{\text{sos}} \, d\nu}{\gamma} = \beta$$

which holds since ν is a probability measure and p_{sos} is non-negative. Lemma 4 justifies the following algorithm, which extracts from a grid the values at which the polynomial p_{sos} is small:

Input moment matrix $M_d(\mathbf{z})$ of measure $\nu = \lambda_{\mathbf{T}} \lambda_{\mathbf{X}} \delta_{y(t,x)}$, small real $\epsilon > 0$, grid points $(t_i, x_j, y_k)_{i,j,k=1,\dots,N} \subset \mathbf{T} \times \mathbf{X} \times \mathbf{Y}$;

Step 1 Compute spectral decomposition (59) of $M_d(\mathbf{z})$ and construct SOS polynomial p_{sos} in (60) with the largest number of terms r such that $\sum_{i=1}^{r} e_i < \epsilon$;

Step 2 For each $i, j, k = 1, \dots, N$, evaluate $p_{i,j,k} := p_{\text{sos}}(t_i, x_j, y_k)$;

Step 3 For each $i, j = 1, \dots, N$, let $y_{i,j} := y_{k^*}$ where $k^* := \arg \min_k p_{i,j,k}$;

Output Approximation $(y_{i,j})_{i,j=1,\dots,N} \subset \mathbf{Y}$ of $y(t,x)$ at grid points $(t_i, x_j)_{i,j=1,\dots,N} \subset \mathbf{T} \times \mathbf{X}$.

The computational burden is modest: an eigenstructure decomposition at Step 1, and grid point evaluations of polynomial p_{sos} at Step 2.

4 The Riemann problem for the Burgers equation

For a numerical illustration, we consider the classical Riemann problem (see e.g., [11]) for a Burgers equation. In particular, we choose the flux

$$f(y) = \frac{1}{4}y^2.$$

The Riemann problem to this conservation law is a Cauchy problem with the following initial condition, piecewise constant with one point of discontinuity:

$$y_0(x) = \begin{cases} l & \text{if } x < 0, \\ r & \text{if } x > 0, \end{cases}$$

where $l, r \in \mathbb{R}$. The solution to the Riemann problem depends strongly on the values of l and r. In particular:

1. If $l > r$, the shock at the initial condition spreads along the characteristics.

2. If $l < r$, the solution is not necessarily unique. The entropy condition allows to select the right solution, which is known as a rarefaction wave.

Both cases are interesting from a numerical point of view for their own reasons. In general, the first case is difficult to tackle because of the discontinuity. In general, numerical schemes based on discretization tend to smoothen out the shock. Indeed, recovering numerically the exact point of discontinuitiy is a challenge for these schemes.

In the second case the solution is continuous, but not necessarily unique. For the Burgers equation, it has been shown that one single entropy condition is sufficient to guarantee uniqueness of the solution [7]. To the best of our knowledge, there is no

similar result for the uniqueness of entropy mv solutions for Burgers equation with concentrated initial data, except for classical solutions [8].

We present numerical results for both cases. We are going to consider $l, r \in \{0, 1\}$. Following the discussion yielding (11), we can assume that the solution takes values only in $\mathbf{Y} = [0, 1]$. The time-space-window on which we consider the solution is $\mathbf{T} = [0, 1]$ and $\mathbf{X} = [L, R] = [-\frac{1}{2}, \frac{1}{2}]$.

Further note that, from the initial condition, we can derive that

$$y(t, L) = l, \quad \forall t \in \mathbf{T}.$$

Moreover, due to positivity of y, the solution on $\mathbf{T} \times \mathbf{X}$ does not depend on the initial condition for $x > \frac{1}{2}$.

Remark on the significance of the numerical results upfront We need to emphasize that these experiments are by no means conclusive. Our implementation is based on the Matlab interface Gloptipoly3 [21] and the SDP solver of MOSEK [5]. The purpose of the numerical examples is to show that our framework actually works in practice and with a proper implementation might actually provide an alternative to schemes based on discretization.

4.1   Shock waves

Let $l = 1$ and $r = 0$. As it has been noticed before, with such an initial condition the solution is discontinuous, for all $t > 0$. The unique analytical solution corresponding to this initial condition is

$$y^*(t, x) = \begin{cases} 1 & x > \frac{t}{4}, \\ 0 & x < \frac{t}{4}. \end{cases} \tag{61}$$

As an objective function, we choose the default implemented in Gloptipoly, which minimizes the trace of the moment matrix. Since the trace is the convex envelope of the rank on the set of matrices with norm less than one, this is likely to cause early convergence of the moment-SOS hierarchy: low rank solutions correspond to measures supported on sets of zero Lebesgue measure. As in this case the marginal of ν with respect to y is supported on $\{0, 1\}$ we expect this criterion to be appropriate to accelerate convergence. Indeed, for $d = 6$ (i.e. moments of degree up to 12) we end up with the following moments for y:

$$(\mathbf{z}_{0,0,k})_{k=0,1,\dots} = (1.0000, 0.6250, 0.6250, 0.6250, 0.6250, \dots)$$

which correspond (up to numerical accuracy) exactly with the moments of the analytic solution.

Localizing the shock In order to approximate the solution from our approximated moments we follow the path lined out in Section 3.3. Applying our algorithm with $\epsilon = 10^{-6}$ yields a polynomial p_{sos} with $r = 54$ terms in the approximate kernel of the moment matrix of size 84, and to the approximated solution represented on Figure 1.

As already mentioned the computed moments can be used in order to approximate the location of the shock at some given time t. Here we will take $t = 0.75$, consequently the shock is located at exactly $x = 0.1875$. We used a standard Godunov scheme (we
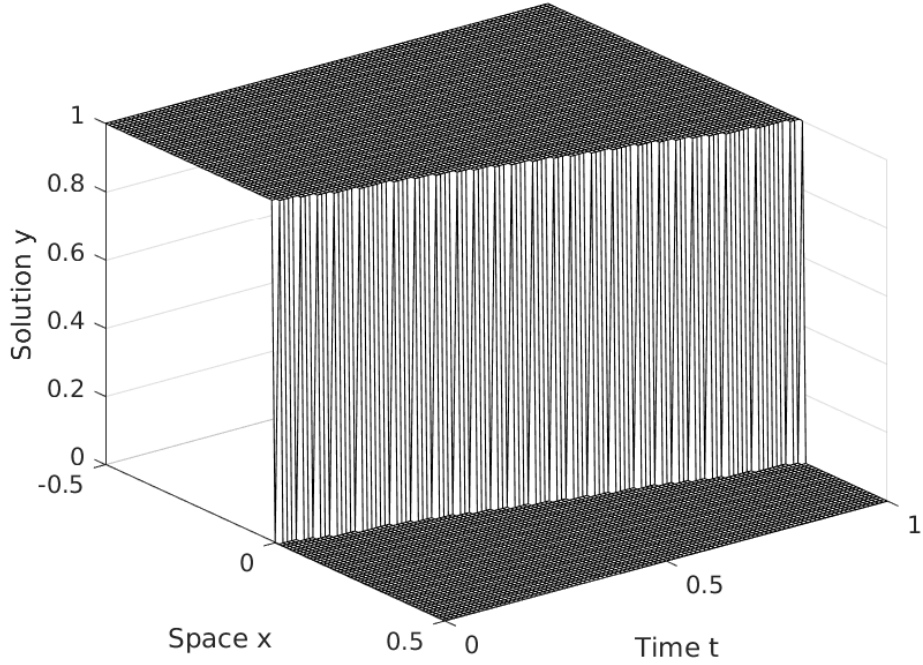
Figure 1: Approximation of the solution $y(t, x)$ obtained with our GMP approach, in the case of a shock.

x	0.1850	0.1855	0.1860	0.1865	0.1870	0.1875	0.1880	0.1885
Godunov	0.9999	0.9991	0.9936	0.9580	0.7647	0.2724	0.0123	0.0000
GMP	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	0.0000	0.0000

Table 1: Approximation of $y(0.75, x)$ with Godunov and GMP.

refer to [27] for more details) to compute the solution up to this time. For space discretization, we took a mesh size of 0.0005 and a consistent discretization in time such that the scheme stays stable. In Table 1, we display the obtained values from this approach on an interval around the shock. We can see the typical behaviour of shock smoothing. In contrast, the values obtained by our GMP approach exactly represent the position of the shock.

4.2 Rarefaction waves

Now let $l = 0$ and $r = 1$. As it has been noticed before, with such an initial condition, entropy conditions are crucial to select the right solution, i.e., the solution with a good physical meaning. The analytical entropy solution corresponding to this example is

$$y(t, x) = \begin{cases} 0 & x \leq 0, \\ \frac{2x}{t} & 0 \leq x \leq \frac{t}{2}, \\ 1 & x \geq \frac{t}{2}. \end{cases} \tag{62}$$
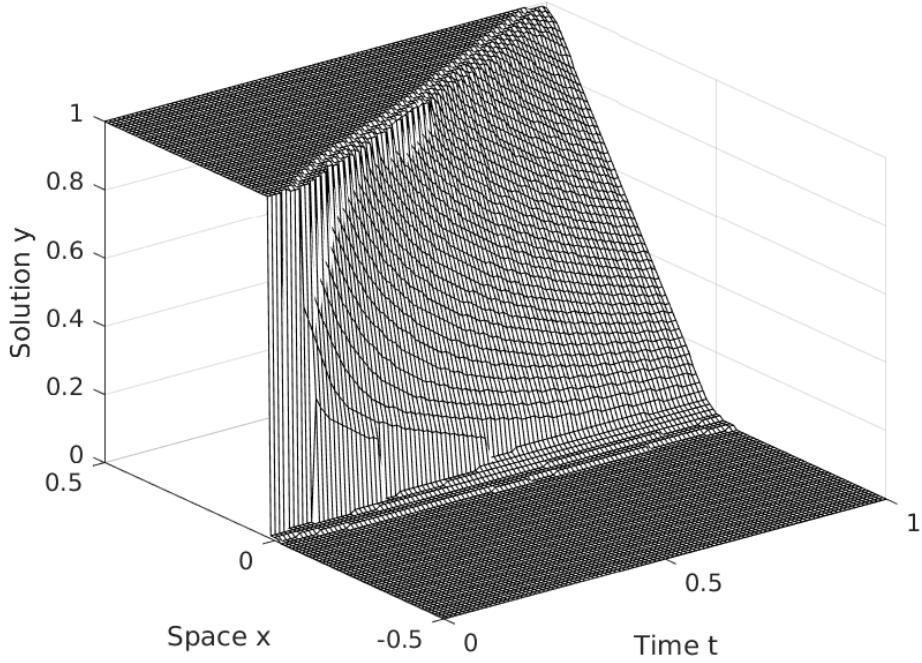
Figure 2: Approximation of the solution $y(t, x)$ obtained with our GMP approach, in the case of a rarefaction wave.

Numerically implementing all entropy pairs of Kruzkhov is possible (as seen in Section 3.1), but heavy. It is known that the entropy $\eta(y) = y^2$ provides all necessary information to make the entropy solution unique for Burgers equation [7]. Then, instead of using all Kruzkhov pairs, we propose the following family of entropies in this example:

$$\eta_k(y) = y^k, \qquad \forall k \in \mathbb{N} \tag{63}$$

and the corresponding polynomial functions q_k. Note that η is strictly convex on $\mathbf{Y} = [0, 1]$. In particular, we do not have to split the measures in (16) into two measures, since there is no absolute value appearing in (63). It is neither necessary to introduce a lifting variable as was discussed in Section 3.1. Finally, we define the sum over all entropy constraints as an objective function to be maximized.

Solving the relaxation of order $d = 6$ (i.e. moments of degree up to 12), we obtain the following moments for the marginal on y:

$$(\mathbf{z}_{0,0,k})_{k=0,1,\ldots} = (1.0000, 0.3750, 0.3333, 0.3125, 0.3000, 0.2917, 0.2857, 0.2812, \ldots)$$

which, again up to numerical accuracy, coincide with the moments of the actual analytic entropy solution. Applying the algorithm from Section 3.3 with $\epsilon = 10^{-6}$ yields a polynomial p_{sos} with $r = 48$ terms in the approximate kernel of the moment matrix of size 84, and the approximated solution represented on Figure 2.

22

5 Conclusion

In this paper, we have provided a new method to solve scalar polynomial hyperbolic partial differential equations. This method relies on the moment-SOS hierarchy surveyed in [24]. More precisely, we have proved that the truncated moments associated to the measure-valued solution formulation converge to the Dirac measure concentrated on entropy solution to the scalar polynomial solution.

The idea of solving linear problems on measures to solve nonlinear differential equation is not new. In the context of nonlinear ordinary differential equations (ODEs), the linear problems involved measures called occupation measures. Roughly speaking, occupation measures allow to measure the time spent by a graph of the trajectory of the ODE in a given subset of the state space. Provided that the nonlinearities considered are polynomial, one can transform the nonlinear ODE into a linear moment problem, in turn solved numerically with the moment-SOS hierarchy, see [25] and the survey [31]. Therefore, the current paper can be seen as an extension to (uncontrolled) PDEs of the results provided in [25] for (controlled) ODEs.

This opens many further research lines. For example:

- One of the most interesting aspect of the notion of very weak solution is the linear formulation on measures of nonlinear differential equations. Such formulations have been useful to solve many problems appearing in the ODE framework, such as optimal control [25] or approximation of region of attraction [20]. The challenge was to prove that the measure formulation was not a relaxation of the original nonlinear problem. For the hyperbolic conservation law studied in our paper, we have used entropy inequalities for that purpose. We are wondering whether it is possible to extend these techniques to the case of other nonlinear PDEs.

- A class of other nonlinear PDEs could be parabolic ones. One of the interest of these equations is that they regularize the solution, whatever is the initial condition. Therefore, as it is done for ODEs in [25], it might be possible to define test functions depending on the solution to the parabolic equation and then define an occupation measure associated to the latter. This together with the relaxed control theory surveyed in [12] might be instrumental to solve optimal control problem for nonlinear parabolic equations.

- The Burgers equation is irreversible. Roughly speaking, given a terminal condition $y(T, x)$ with $T > 0$, there exists a continuum of initial conditions yielding $y(T, x)$, see e.g., [18]. Such a continuum can be described with measures and, hence, our linear formulation might be useful to solve such inverse problems, extending to PDEs what was developed in [20] for ODEs.

A   Proof of Theorem 1

Proof The proof is divided into two steps. The first step consists in proving that v can be replaced by $y(t, x)$ or $\langle \mu_{(t,x)}, y \rangle$. The second step aims at proving the contraction inequality given in (7). In each step, a special choice of test function is done in order to prove the result.

• First step: Doubling variable

Let us consider the entropy pair given in (8). For all $(s, z) \in \mathbb{R}_+ \times \mathbb{R}$, we choose $v = y(s, z)$, where y is an entropy solution to (12):

$$
\int_0^T \int_{\mathbb{R}} \left(\frac{\partial \psi}{\partial t} \langle \mu_{(t,x)}, |y - y(s, z)| \rangle + \frac{\partial \psi}{\partial x} \langle \mu_{(t,x)}, \operatorname{sign}(y - y(s, z))(f(y) - f(y(s, z))) \rangle \right) dx\, dt
$$
$$
+ \int_{\mathbb{R}} \psi(0, x) \langle \sigma_0, |y - y(s, z)| \rangle dx - \int_{\mathbb{R}} \psi(T, x) \langle \mu_{(t,x)}, |y - y(s, z)| \rangle dx.
$$
(64)

Similarly, for all $(s, z) \in \mathbb{R}_+ \times \mathbb{R}$, we set $v = \langle \mu_{(t,x)}, y \rangle$ in (3) and use the fact that μ is a probability measure:

$$
\int_0^T \int_{\mathbb{R}} \left(\frac{\partial \psi}{\partial s} \langle \mu_{(t,x)}, |y - y(s, z)| \rangle + \frac{\partial \psi}{\partial z} \langle \mu_{(t,x)}, \operatorname{sign}(y - y(s, z))(f(y) - y(s, z)) \rangle \right) ds\, dz
$$
$$
+ \int_{\mathbb{R}} \psi(0, z) \langle \mu_{(t,x)}, |y - y_0(z)| \rangle dz \geq 0.
$$
(65)

Let us choose $\tilde{\psi} := \tilde{\psi}(t, x, s, z)$. Thanks to the two latter inequalities, one has

$$
\int_{\mathbb{R}_+} \int_{\mathbb{R}} \int_{\mathbb{R}_+} \int_{\mathbb{R}} \left(\left(\frac{\partial \psi}{\partial s} + \frac{\partial \psi}{\partial t} \right) \langle \mu_{(t,x)}, |y - y(s, z)| \rangle \right.
$$
$$
\left. + \left(\frac{\partial \psi}{\partial z} + \frac{\partial \psi}{\partial x} \right) \langle \mu_{(t,x)}, \operatorname{sign}(y - y(s, z))(f(y) - f(y(s, z))) \rangle \right) ds\, dz\, dx\, dt
$$
$$
+ \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(t, x, 0, z) \langle \mu_{(t,x)}, |y - y_0(z)| \rangle dz\, dx
$$
$$
+ \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(0, x, s, z) \langle \sigma_0, |y - y(s, z)| \rangle dz\, dx \geq 0.
$$
(66)

Let $\rho_\varepsilon \in \mathscr{C}^1(\mathbb{R})$ satisfy

$$
\int_{\mathbb{R}} \rho_\varepsilon(s) ds = 1.
$$
(67)

For all $\varphi \in \mathscr{C}_c^1(\mathbb{R}_+ \times \mathbb{R})$, one defines $\tilde{\psi}$ as follows

$$
\tilde{\psi}(t, x, s, z) = \frac{1}{\varepsilon^2} \rho_\varepsilon \left(\frac{t - s}{2\varepsilon} \right) \rho_\varepsilon \left(\frac{x - z}{2\varepsilon} \right) \varphi \left(\frac{t + s}{2}, \frac{x + z}{2} \right).
$$
(68)

Therefore

$$
\left(\frac{\partial \tilde{\psi}}{\partial t} \tilde{\psi} + \frac{\partial \tilde{\psi}}{\partial s} \right) = \frac{1}{\varepsilon^2} \frac{\partial \varphi}{\partial t} \left(\frac{t + s}{2}, \frac{x + z}{2} \right) \rho_\varepsilon \left(\frac{x - z}{2\varepsilon} \right) \rho_\varepsilon \left(\frac{t - s}{2\varepsilon} \right)
$$
(69)

and

$$\left(\frac{\partial \tilde{\psi}}{\partial x} + \frac{\partial \tilde{\psi}}{\partial z}\right) = \frac{1}{\varepsilon^2}\frac{\partial \varphi}{\partial x}\left(\frac{t+s}{2}, \frac{x+z}{2}\right))\rho_\varepsilon\left(\frac{x-z}{2\varepsilon}\right)\rho_\varepsilon\left(\frac{t-s}{2\varepsilon}\right). \tag{70}$$

We aim at proving that

$$\rho_\varepsilon\left(\frac{t-s}{\varepsilon}\right) \to \delta(t-s) \text{ as } \varepsilon \to 0, \tag{71}$$

and

$$\rho_\varepsilon\left(\frac{x-z}{\varepsilon}\right) \to \delta(x-z) \text{ as } \varepsilon \to 0, \tag{72}$$

so that we will have, thanks to (69) and (70)

$$\int_{\mathbb{R}_+}\int_\mathbb{R}\left(\frac{\partial \varphi}{\partial t}\langle\mu_{(t,x)}, |y-y(t,x)|\rangle + \frac{\partial \varphi}{\partial x}\langle\mu_{(t,x)}, \text{sign}(y-y(t,x))(f(y)-f(y(t,x)))\rangle\right)dx\,dt$$
$$+ \int_\mathbb{R}\varphi(0,x)\langle\mu_{(t,x)}, |y-y_0(x)|\rangle dx \geq 0. \tag{73}$$

Noticing that, for a fixed $t > 0$,

$$2\int_\mathbb{R}\frac{1}{2\varepsilon}\rho\left(\frac{t-s}{\varepsilon}\right)ds = 2, \tag{74}$$

then, up to a change of variable, one has, for any continuous function $\phi := \phi(t,s)$

$$\int_\mathbb{R}\int_\mathbb{R}\frac{1}{\varepsilon}\rho\left(\frac{t-s}{\varepsilon}\right)\phi(t,s)dt\,ds = 2\int_\mathbb{R}\phi(t,t)dt. \tag{75}$$

Since ϕ is continuous and lies in a compact set, then it is uniformly continuous. Therefore, one has

$$|t-s| \leq 2\varepsilon \Rightarrow |\phi(t,s) - \phi(t,t)| \leq M_\varepsilon. \tag{76}$$

Therefore, for any positive value k, it follows that

$$\left|\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\frac{1}{\varepsilon}\rho\left(\frac{t-s}{2\varepsilon}\right)\phi(t,s) - \phi(t,t)dt\,ds\right| \leq M_\varepsilon\int_\mathbb{R}\int_{|t-s|\leq 2\varepsilon, |s|\leq k, |t|\leq k}\frac{1}{\varepsilon}\rho\left(\frac{t-s}{2\varepsilon}\right)dt\,ds$$
$$\leq 4M_\varepsilon k. \tag{77}$$

Note that

$$M_\varepsilon \to 0 \text{ as } \varepsilon \to 0. \tag{78}$$

Hence,

$$\left|\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\frac{1}{\varepsilon}\rho\left(\frac{t-s}{2\varepsilon}\right)(\phi(t,s) - \phi(t,t))dt\,ds\right| \to 0 \text{ as } \varepsilon \to 0. \tag{79}$$

Similarly, one can prove that, for any continuous function $\tilde{\phi} := \tilde{\phi}(x,z)$

$$\left|\int_\mathbb{R}\int_\mathbb{R}\frac{1}{\varepsilon}\rho\left(\frac{x-z}{2\varepsilon}\right)(\tilde{\phi}(x,z) - \tilde{\phi}(x,x))dt\,ds\right| \to 0 \text{ as } \varepsilon \to 0. \tag{80}$$

Finally, using (69) and (70), the equation (66) converges to

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} \left(\frac{\partial \varphi}{\partial t} \langle \mu_{(t,x)}, |y - y(t,x)| \rangle + \frac{\partial \varphi}{\partial x} \langle \mu_{(t,x)}, \mathrm{sign}(y - y(t,x))(f(y) - f(y(t,x))) \rangle \right) dx \, dt$$

$$+ \int_{\mathbb{R}} \varphi(0,x) \langle \sigma_0, |y - y_0(x)| \rangle dx \geq 0,$$

(81)

as ε goes to 0.

• Second step: Contraction inequality

Given two positive values $r \in \mathbb{R}$ and $T \in \mathbb{R}_+$, let us choose φ as follows:

$$\varphi(t,x) = \theta_\varepsilon(t)\Delta_\varepsilon(t,x), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}$$

(82)

where

$$\theta_\varepsilon(t) := \begin{cases} 1 & t \leq T, \\ \dfrac{T-t}{\varepsilon} & T \leq t \leq T + \varepsilon, \\ 0 & t \geq T + \varepsilon, \end{cases}$$

(83)

and

$$\Delta_\varepsilon(t,x) := \begin{cases} 1 & |x| \leq r + C(T-t), \\ \dfrac{r + C(T-t) - |x|}{\varepsilon} & r + C(T-t) \leq |x| \leq r + C(T-t) + \varepsilon, \\ 0 & |x| \geq r + C(T-t) + \varepsilon \end{cases}$$

(84)

where C denotes the Lipschitz constant of the function f. Differentating θ_ε with respect to t yields

$$\frac{\partial \theta_\varepsilon(t)}{\partial t} := \begin{cases} -\dfrac{1}{\varepsilon} & T \leq t \leq T + \varepsilon, \\ 0 \end{cases}$$

(85)

Differentiating Δ_ε with respect to t yields

$$\frac{\partial \Delta_\varepsilon}{\partial t} := \begin{cases} -\dfrac{C}{\varepsilon} & r + C(T-t) \leq |x| \leq r + C(T-t) + \varepsilon, \\ 0 & |x| \geq r + C(T-t) + \varepsilon, \end{cases}$$

(86)

Finally, differentiating Δ_ε with respect to x yields

$$\frac{\partial \Delta_\varepsilon}{\partial x} := \begin{cases} -\dfrac{\mathrm{sign}(x)}{\varepsilon} & r + C(T-t) \leq |x| \leq r + C(T-t) + \varepsilon, \\ 0 & |x| \geq r + C(T-t) + \varepsilon. \end{cases}$$

(87)

26

Hence, using these functions in (81), one has

$$-\frac{1}{\varepsilon}\int_T^{T+\varepsilon}\int_{\mathbb{R}}\Delta_\varepsilon(t,x)\langle\mu_{(t,x)},|y-y(t,x)|\rangle dt\,dx$$

$$-\frac{L}{\varepsilon}\int_0^T\int_{r+C(T-t)\le|x|\le r+C(T-t)+\varepsilon}\langle\mu_{(t,x)},|y-y(t,x)|\rangle dt\,dx$$

$$-\frac{1}{\varepsilon}\int_{\mathbb{R}_+}\int_{r+C(T-t)\le|x|\le r+C(T-t)+\varepsilon}\operatorname{sign}(x)\langle\mu_{(t,x)},\operatorname{sign}(y-y(t,x))(f(y)-f(y(t,x)))\rangle dt\,dx$$

$$+\int_{\mathbb{R}}\varphi(0,x)\langle\sigma_0,|y-y_0(x)|\rangle dx\ge 0.$$

$$(88)$$

Note that one has

$$\operatorname{sign}(x)\langle\mu_{(t,x)},\operatorname{sign}(y-y(t,x))(f(y)-f(y(t,x)))\rangle\le C\langle\mu_{(t,x)},|y-y(t,x)|\rangle.\qquad(89)$$

Therefore, (88) becomes

$$\frac{1}{\varepsilon}\int_T^{T+\varepsilon}\int_{\mathbb{R}}\Delta_\varepsilon(t,x)\langle\mu_{(t,x)},|y-y(t,x)|\rangle dt\,dx\le\int_{\mathbb{R}}\varphi(0,x)\langle\sigma_0,|y-y_0(x)|\rangle dx.$$

From this latter equation, one can conclude the proof when ε goes to 0. Indeed, the left hand side of the inequality can be bounded as follows:

$$\frac{1}{\varepsilon}\left|\int_T^{T+\varepsilon}\int_{\mathbb{R}}\Delta_\varepsilon(t,x)\langle\mu_{(t,x)},|y-y(t,x)|\rangle dt\,dx\right|$$

$$\le\frac{1}{\varepsilon}\left|\int_T^{T+\varepsilon}\int_{|x|\le r+\varepsilon}\frac{r+C(T-t)-|x|}{\varepsilon}\langle\mu_{(t,x)},|y-y(t,x)|\rangle dt\,dx\right|\qquad(90)$$

$$\le\frac{1}{\varepsilon}\int_T^{T+\varepsilon}\int_{|x|\le r+\varepsilon}\left|\frac{r+C(T-t)-|x|}{\varepsilon}\right|\left|\langle\mu_{(t,x)},|y-y(t,x)|\rangle\right| dt\,dx.$$

Noticing that, for any $t\in[T,T+\varepsilon]$ and any $|x|\le r+\varepsilon$, one has

$$\left|\frac{r+C(T-t)-|x|}{\varepsilon}\right|\le 1,\qquad(91)$$

then one has

$$\frac{1}{\varepsilon}\int_T^{T+\varepsilon}\int_{\mathbb{R}}\Delta_\varepsilon(t,x)\langle\mu_{(t,x)},|y-y(t,x)|\rangle dt\,dx\le\frac{1}{\varepsilon}\int_T^{T+\varepsilon}\int_{|x|\le r+\varepsilon}\left|\langle\mu_{(t,x)},|y-y(,x)|\rangle\right| dt\,dx.$$

$$(92)$$

Moreover, one has

$$\frac{1}{\varepsilon}\int_T^{T+\varepsilon}\int_{|x|\le r+\varepsilon}\left|\langle\mu_{(t,x)},|y-y(,x)|\rangle\right| dt\,dx\to\int_{|x|\le r}\langle\sigma_T,|y-y(T,x)|\rangle dx\ \text{as}\ \varepsilon\to 0.$$

$$(93)$$

Similarly, one can prove that

$$\int_{\mathbb{R}}\varphi(0,x)\langle\sigma_0,|y-y_0(x)|\rangle dx\to\int_{|x|\le r+CT}\langle\sigma_0,|y-y_0(x)|\rangle dx\ \text{as}\ \varepsilon\to 0.\qquad(94)$$

Finally, it yields, for all positive values T and r

$$\int_{|x|\leq r} \langle \sigma_T, |y - y(T,x)|\rangle dx \leq \int_{|x|\leq r+CT} \langle \sigma_0, |y - y_0(x)|\rangle dx. \tag{95}$$

This concludes the proof of Theorem 1. $\qquad\square$

References

[1] Y. Brenier. Solution by convex minimization of the Cauchy problem for hyperbolic systems of conservation laws with convex entropy. `arXiv:1710.03754`, 2017.

[2] H. Brezis. Functional analysis, Sobolev spaces and partial differential equations. Springer, 2010.

[3] S. I. Chernyshenko, P. Goulart, D. Huang, and A. Papachristodoulou. Polynomial sum of squares in fluid dynamics: a review with a look ahead. Phil. Trans. R. Soc. A, 372(2020):20130350, 2014.

[4] M. Claeys, R. Sepulchre. Reconstructing trajectories from the moments of occupation measures. Proc. IEEE Conf. on Decision and Control, 2014.

[5] J. Dahl. Extending the conic optimizer in MOSEK with semidefinite cones. Proc. Intl. Symp. Math. Prog., Berlin, 2012.

[6] C. M. Dafermos. Hyperbolic conservation laws in continuum physics. Springer, 2006.

[7] C. DeLellis, F. Otto, M. Westdickenberg. Minimal entropy conditions for Burgers equation. Quarterly of applied mathematics 62(4):687–700, 2004.

[8] S. Demoulini, D. M. A. Stuart, A. E. Tzavaras. Weak-strong uniqueness of dissipative measure-valued solutions for polyconvex elastodynamics. Archive for Rational Mechanics and Analysis 205(3):927–961, 2012.

[9] B. Després, F. Lagoutière. Contact discontinuity capturing schemes for linear advection and compressible gas dynamics. Journal of Scientific Computing 16(4):479–524, 2001.

[10] R.J. DiPerna. Measure-valued solutions to conservation laws. Archive for Rational Mechanics and Analysis 88(3):223–270, 1985.

[11] L.C. Evans. Partial differential equations. American Mathematical Society, 2010.

[12] H.O. Fattorini. Infinite dimensional optimization and control theory. Cambridge University Press, 1999.

[13] E. Feireisl, M. Lukáčová-Medvid'ová, H. Mizerová. Convergence of finite volume schemes for the Euler equations via dissipative measure-valued solutions. `arXiv:1803.08401`, 2018.

[14] U. S. Fjordholm, S. Mishra, E. Tadmor. On the computation of measured-valued solutions. Acta Numerica 25:567-679, 2016.

[15] U. S. Fjordholm, R. Käppeli, S. Mishra, E.Tadmor. Construction of approximate entropy measure-valued solutions for hyperbolic systems of conservation laws. Foundations of Computational Mathematics 17(3):763–827, 2017.

[16] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. Matematicheskii Sbornik 89(3):271–306, 1959.

[17] D. Goluskin, G. Fantuzzi. Bounds on mean energy in the Kuramoto-Sivashinsky equation computed using semidefinite programming. `arXiv:1802.08240`, 2018.

[18] L. Gosse, E. Zuazua. Filtered gradient algorithms for inverse design problems of one-dimensional Burgers equation. Pages 197-227 in L. Gosse, R. Natalini (Editors). Innovative algorithms and analysis. SINDAM Series, Springer, 2017.

[19] D. Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. Pacific J. Math. 132(1):35–62, 1988.

[20] D. Henrion, M. Korda. Convex computation of the region of attraction of polynomial control systems. IEEE Trans. Autom. Control 59(2):297–312, 2014.

[21] D. Henrion, J. B. Lasserre, J. Löfberg. Gloptipoly 3: moments, optimization and semidefinite programming. Optimization Methods & Software 24(4-5):761–779, 2009.

[22] M. Korda, D. Henrion, J. B. Lasserre. Moments and convex optimization for analysis and control of nonlinear partial differential equations. `arXiv:1804.07565`, 2018.

[23] S. N. Kružkov. First order quasilinear equations in several independent variables. Mathematics of the USSR-Sbornik 10(2):217, 1970.

[24] J. B. Lasserre. Moments, positive polynomials and their applications. Imperial College Press, 2010.

[25] J. B. Lasserre, D. Henrion, C. Prieur, E. Trélat. Nonlinear optimal control via occupation measures and LMI relaxations. SIAM Journal on Control and Optimization 47(4):1643–1666, 2008.

[26] P. Lax. Shock waves and entropy. Pages 603–634 in E. H. Zarantonello (Editor). Contributions to nonlinear functional analysis. Elsevier, 1971.

[27] R. J. LeVeque. Numerical methods for conservation laws. Lectures in Mathematics, ETH Zürich, 1992.

[28] M. Mevissen, J. B. Lasserre, D. Henrion. Moment and SDP relaxation techniques for smooth approximations of problems involving nonlinear differential equations. Proc. IFAC World Congress on Automatic Control, 2011.

[29] J. Málek, J. Nečas, M. Rokyta, M. Růžička. Weak and measure-valued solutions to evolutionary PDEs. CRC Press, 1996.

[30] E. Y. Panov. Uniqueness of the solution of the Cauchy problem for a first order quasilinear equation with one admissible strictly convex entropy. Mathematical Notes 55(5):517–525, 1994.

[31] D. Henrion, E. Pauwels. Linear conic optimization for nonlinear optimal control. Pages 121-134 in S. Ahmed, M. Anjos, T. Terlaky (Editors). Advances and Trends in Optimization with Engineering Applications. SIAM, 2017.

[32] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana University Mathematics Journal 42(3):969–984, 1993.

[33] J. Rubio. The global control of shock waves. Pages 355-369 in M. Grosser, G. Hörmann, M. Kunzinger and L. Oberguggenberger (Editors). Nonlinear theory of generalized functions. Erwin Schrödinger Institute, Vienna, 1997.

[34] G. B. Whitham. Linear and nonlinear waves. John Wiley & Sons, 2011.