A moment approach for entropy solutions of parameter-dependent hyperbolic conservation laws

Clément Cardoen, Swann Marx, Anthony Nouy, Nicolas Seguin

A moment approach for entropy solutions of parameter-dependent hyperbolic conservation laws

Clément Cardoen[1], Swann Marx[2], Anthony Nouy[1], and Nicolas Seguin[3]

[1]Nantes Université, Centrale Nantes, Laboratoire de Mathématiques Jean Leray, CNRS UMR 6629, France
[2]Nantes Université, Centrale Nantes, LS2N, CNRS UMR 6004, France
[3]Inria, Centre de l'Université Côte d'Azur, antenne de Montpellier, Imag, UMR CNRS 5149, Université de Montpellier, France

Abstract

We propose a numerical method to solve parameter-dependent hyperbolic partial differential equations (PDEs) with a moment approach, based on a previous work from Marx et al. (2020). This approach relies on a very weak notion of solution of nonlinear equations, namely parametric entropy measure-valued (MV) solutions, satisfying linear equations in the space of Borel measures. The infinite-dimensional linear problem is approximated by a hierarchy of convex, finite-dimensional, semidefinite programming problems, called Lasserre's hierarchy. This gives us a sequence of approximations of the moments of the occupation measure associated with the parametric entropy MV solution, which is proved to converge. In the end, several post-treatments can be performed from this approximate moments sequence. In particular, the graph of the solution can be reconstructed from an optimization of the Christoffel-Darboux kernel associated with the approximate measure, that is a powerful approximation tool able to capture a large class of irregular functions. Also, for uncertainty quantification problems, several quantities of interest can be estimated, sometimes directly such as the expectation of smooth functionals of the solutions. The performance of our approach is evaluated through numerical experiments on the inviscid Burgers equation with parametrised initial conditions or parametrised flux function.

[1]{clement.cardoen,anthony.nouy}@ec-nantes.fr
[2]swann.marx@ls2n.fr
[3]nicolas.seguin@inria.fr

1 Introduction

Non-linear hyperbolic conservation laws model numerous physical phenomena in fluid mechanics, traffic flow or non-linear acoustics [8, 38]. The numerical computation of such equations is often a challenge since their solutions may present discontinuities, even if the initial data are smooth. Numerous numerical methods exist to approximate them, amongst which we may cite finite volume or finite difference schemes [27] or the front-tracking method [19]. We are interested in this paper in solving parameter-dependent hyperbolic conservation laws, which are considered for various tasks in data assimilation [6], uncertainty quantification [1, 34, 5, 39, 3], sensitivity analysis [7], or error analysis [13]. The parameters in our context appear in the initial data and in the flux function and are associated with a probability measure. The computation of approximate solutions for many instances of the parameters is usually prohibitive and require reduced order models.

Model order reduction methods aim at providing an approximation of the solution $u(z, \xi)$, depending on physical variables z and parameters ξ, that can be efficiently evaluated. They either rely on an explicit approximation of the solution map $\xi \mapsto u(\cdot, \xi)$ or an approximation of the solution manifold $\{u(\cdot, \xi) : \xi \in \Xi\}$ by some dimension reduction method. The main challenge for models driven by conservation laws is that the solution maps and solution manifolds are highly nonlinear (in particular due to the presence of discontinuities), that require the introduction of nonlinear approximation or dimension reduction methods. Several model reduction methods based on compositions have been proposed, that include methods based on parameter-dependent changes of variables [36, 15] or deep learning methods using neural networks [23]. These methods usually require high computational resources and huge training data for the approximation of highly nonlinear solution maps.

Here, we follow a different approach and propose a new surrogate modelling method. It is an extension of [29] to parameter-dependent or random conservation laws. Whereas it is classical to seek entropy weak solutions to hyperbolic conservation laws [8, 22], we are rather interested in so-called entropy measure-valued (MV) solutions, an even weaker notion of solution, introduced by DiPerna in [10, 32]. To a MV solution corresponds an occupation measure, whose marginal is the MV solution. Even if this notion of solution is very weak, there is a correspondence with entropy weak solution. The measure concentrated on the graph of the entropy weak solution is a MV solution. It is worth noting that the formulation in the setting of MV solutions leads to a linear problem, so that some efficient tools from convex analysis can be applied.

We start with a theoretical framework for parameter-dependent conservation laws similar to the one of [31, 30]. However, in our case, we introduce a weak-parametric formulation of the problem, where the classical entropy weak formulation is also integrated with respect to the parameter. The purpose of this formulation is to obtain an equivalent definition of parameter-dependent entropy MV solutions using the moments of the associated occupation measure with respect to all the variables, including the parameters. Under the assumption that flux function is polynomial and that the initial data can be described by semi-algebraic functions, the entropy formulation becomes a set of linear constraints on the moments of the occupation measure and we can follow the procedure initiated in [29]. Indeed, this allows us to consider the problem

as a generalized moment problem (GMP), an infinite-dimensional optimization problem over sequences of moments of measures, where both the cost and the constraints are linear with respect to the moments of the measures. Powerful results from real algebraic geometry allow to reformulate the constraint that a sequence is a moment sequence into tractable semi-definite constraints. This problem is then solved using Lasserre's (moment sum-of-squares) hierarchy [24], which consists in solving a sequence of convex semi-definite programs of increasing size to approximate the moments of the occupation measure. Note that the use of Lasserre's hierarchy for solving PDEs has been also recently considered in [16], although with a different approach where the considered measure is defined on an infinite-dimensional function space, and assumed to be concentrated on the solution of the PDE.

Obtaining an approximation of the moments can be costly, but once this offline computation is performed, efficient online post-treatments are possible. First, we can naturally obtain expectations of variables of interest that are functions of the moments of the solution. Also, the graph of the entropy weak solution (for any parameter value) can be recovered using a localization property of the Christoffel-Darboux kernel of the approximate occupation measure, following the methodology proposed in [28]. This powerful approximation method allows to capture efficiently discontinuities in the solutions. Using the moment completion technique from [18], one can also have access to other quantities of interest, such as statistical moments of point-wise evaluations of the solution.

Outline This paper is organized as follows. We first introduce some notations and the problem considered. Section 2 introduces different notions of solutions for parametrised scalar conservation laws and examines some links between these notions. Section 3 introduces the moment-SOS hierarchy and indicates how to perform several post-treatments such as retrieving the graph of the solution or estimating statistical moments of the solution. Finally, Section 4 presents some numerical experiments.

1.1 Notations

For $\mathcal{X} \subset \mathbb{R}^n$, with $n \in \mathbb{N}$, let $\mathcal{C}(\mathcal{X})$, $\mathcal{C}_0(\mathcal{X})$ and $\mathcal{C}^1_c(\mathcal{X})$ denote the space of functions on $\mathcal{X}$ that are continuous, continuous and vanishing at infinity and continuously differentiable with compact support, respectively. The sets of signed Borel measures and positive Borel measures are denoted $\mathcal{M}(\mathcal{X})$ and $\mathcal{M}(\mathcal{X})_+$, respectively. The set of probability measures on $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$. The measure $\lambda_{\mathcal{X}} \in \mathcal{M}(\mathcal{X})_+$ denotes the Lebesgue measure on $\mathcal{X}$, and for $B \subset \mathcal{X}$ a Borel set, $|B|$ denotes its Lebesgue measure. Given a vector $\mathbf{w} = (w_1, \dots, w_n)$, we denote by $\mathbb{R}[\mathbf{w}]$ the ring of real multivariate polynomials in the variable $w_1, \dots, w_n$, and for a multi-index $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$, $\mathbf{w}^{\boldsymbol{\alpha}} := w_1^{\alpha_1} \dots w_n^{\alpha_n}$. Given a positive Borel measure μ, we denote by $\mathrm{supp}(\mu)$ its support, defined as the smallest closed set whose complement has measure zero.

1.2 Definition of the problem

We consider parameter-dependent scalar hyperbolic conservation laws that are formulated as a Cauchy problem

$$\partial_t u(t,\mathbf{x},\boldsymbol{\xi}) + \operatorname{div}_{\mathbf{x}} \mathbf{f}(u(t,\mathbf{x},\boldsymbol{\xi}),\boldsymbol{\xi}) = 0, \quad (t,\mathbf{x},\boldsymbol{\xi}) \in \mathbb{R}_+ \times \mathbb{R}^n \times \boldsymbol{\Xi}, \tag{1a}$$

$$u(0,\mathbf{x},\boldsymbol{\xi}) = u_0(\mathbf{x},\boldsymbol{\xi}), \quad (\mathbf{x},\boldsymbol{\xi}) \in \mathbb{R}^n \times \boldsymbol{\Xi}, \tag{1b}$$

where $t \in \mathbb{R}_+$ is the time variable, $\mathbf{x} \in \mathbb{R}^n$ is the space variable, and where $\boldsymbol{\xi}$ is a parameter in a parameter set $\boldsymbol{\Xi} \subset \mathbb{R}^p$, $p \in \mathbb{N}$. Then data are the flux $\mathbf{f} : \mathbb{R} \times \boldsymbol{\Xi} \to \mathbb{R}^n$ and the initial condition $u_0 : \mathbb{R}^n \times \boldsymbol{\Xi} \to \mathbb{R}$.

The parameter set $\boldsymbol{\Xi}$ is equipped with a probability measure ρ. We assume that $\boldsymbol{\Xi}$ is compact and the initial condition $u_0 \in \mathcal{L}^\infty(\boldsymbol{\Xi}; \mathcal{L}^\infty(\mathbb{R}^n))$. Moreover, we shall make the following assumptions on $\mathbf{f}$.

Assumption 1.1. *For all $K \subset \mathbb{R}$ compact, there exists $C_K \in \mathbb{R}$ such that for all $u \in K$ and $\rho-$a.e., $\|\mathbf{f}(u,\boldsymbol{\xi})\| \leq C_K$. Moreover, $\rho-$a.e., $\mathbf{f}(\cdot,\boldsymbol{\xi}) \in \mathcal{C}^1(\mathbb{R};\mathbb{R}^n)$.*

This assumption is satisfied for a polynomial function f, that will be assumed in the next section. Then for the sake of simplicity, we restrict our analysis to this setting. A more general setting can be found in [30].

Note that the initial data u_0 and the flux f may depend on distinct parameters but for the sake of clarity, and without loss of generality, we indicate a dependence on the same set of parameters $\boldsymbol{\xi}$.

2 Notions of parameter-dependent solution

2.1 Parametric entropy solution

We start by introducing the notion of parametric entropy weak solution, that is defined point-wise in the parameter domain. This may be considered as a strong-parametric solution, that is a straightforward notion of solution when a parameter is considered, see, e.g., [31] or [30].

Definition 2.1 (Entropy pairs)**.** Let η be a locally Lipschitz and convex function from $\mathbb{R}$ to $\mathbb{R}$. Let $\mathbf{q} : \mathbb{R} \times \boldsymbol{\Xi} \to \mathbb{R}^n$ such that $\partial_u \mathbf{q}(u,\boldsymbol{\xi}) = \eta'(u)\partial_u \mathbf{f}(u,\boldsymbol{\xi})$ for ρ-almost $\boldsymbol{\xi}$ and almost all $u \in \mathbb{R}$. Then $(\eta,\mathbf{q})$ is called an entropy pair associated with conservation law (1a).

We may notice that for an entropy pair $(\eta,\mathbf{q})$, for ρ-almost all $\boldsymbol{\xi}$, $\mathbf{q}(\cdot,\boldsymbol{\xi})$ is a locally Lipschitz function.

We now introduce three specific families of entropy pairs, each of them having a particular theoretical or numerical objective.

Definition 2.2 ($\mathcal{C}^1$ family of entropy pairs)**.** The $\mathcal{C}^1$ family of entropy pairs, denoted $\mathcal{E}_{\mathcal{C}}$, is defined as the set of entropy pairs $(\eta,\mathbf{q})$ such that $\eta \in \mathcal{C}^1(\mathbb{R})$ and for ρ-almost all $\boldsymbol{\xi}$, $\mathbf{q}(\cdot,\boldsymbol{\xi}) \in \mathcal{C}^1(\mathbb{R})$.

Note under assumption 1.1, if $(\eta,\mathbf{q})$ is an entropy pair with $\eta \in \mathcal{C}^1$, then $(\eta,\mathbf{q}) \in \mathcal{E}_{\mathcal{C}}$. The $\mathcal{C}^1$ family of entropy pairs is related to the (opposite of the) thermodynamic entropy and to the second law of thermodynamics, for fluid dynamics models. The conservation law (1a), for a fixed $\boldsymbol{\xi}$, can be seen as a simplification of such models.

Definition 2.3 (Kruzhkov family of entropy pairs)**.** The Kruzhkov family of entropy pairs is defined by

$$\mathcal{E}_K := \{(\eta_v, \mathbf{q}_v) : v \in \mathbb{R}\},$$

where for all $v \in \mathbb{R}$, for all $u \in \mathbb{R}$ and for ρ-almost all $\boldsymbol{\xi}$, $\eta_v(u) := |u - v|$ and $\mathbf{q}_v(u, \boldsymbol{\xi}) := \operatorname{sign}(u - v)(\mathbf{f}(u, \boldsymbol{\xi}) - \mathbf{f}(v, \boldsymbol{\xi}))$.

Compared to $\mathcal{E}_\mathcal{C}$, the family $\mathcal{E}_K$ has the advantage of being explicitly described and of carrying strong results coming from Kruzhkov's fundamental paper [22], allowing to obtain some theoretical results, such as uniqueness and stability.

Definition 2.4 (Polynomial family of entropy pairs)**.** The polynomial family of entropy pairs $\mathcal{E}_P$ is defined as the set of entropy pairs $(\eta, \mathbf{q})$ such that η is a polynomial function. If $\mathbf{f}$ is a polynomial function, then, for ρ-almost all $\boldsymbol{\xi}$, $\mathbf{q}(\cdot, \boldsymbol{\xi})$ is also a polynomial function.

In the case of a uniformly convex flux function, a single polynomial entropy can be sufficient to select the relevant solution (see e.g. [9, 21]). Actually, our motivation is different. In numerical experiments, we shall use subsets of the polynomial family of entropy pairs. The SOS-moment (Lasserre's) hierarchy, later exposed in this paper, relies indeed on a polynomial setting. Although it is possible to implement our numerical method with $\mathcal{E}_K$ as in [29], it is easier to do so with subsets of $\mathcal{E}_P$ when possible.

Definition 2.5 (Parametric entropy solution)**.** Consider a family of entropy pairs $\mathcal{E}$. Let $u_0 : \mathbb{R}^n \times \boldsymbol{\Xi} \to \mathbb{R}$ such that $u_0(\cdot, \boldsymbol{\xi}) \in \mathcal{L}^\infty(\mathbb{R}^n)$ ρ-almost everywhere and $\mathbf{f}$ satisfy Assumption 1.1. A function $u : \mathbb{R}_+ \times \mathbb{R}^n \times \boldsymbol{\Xi} \to \mathbb{R}$ such that for ρ-almost all $\boldsymbol{\xi}$, $u(\cdot, \cdot, \boldsymbol{\xi}) \in \mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$ is a parametric entropy solution for $\mathcal{E}$ if, for all $(\eta, \mathbf{q}) \in \mathcal{E}$, for all non-negative test functions $\psi \in \mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n)$ and ρ-almost all $\boldsymbol{\xi}$, it satisfies

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}^n} \left(\partial_t \psi \eta(u) + \nabla_\mathbf{x} \psi \cdot \mathbf{q}(u, \boldsymbol{\xi})\right) d\mathbf{x}dt + \int_{\mathbb{R}^n} \psi \eta(u_0) d\mathbf{x} \geq 0. \tag{2}$$

Proposition 2.6. *A function u is a parametric entropy solution for $\mathcal{E}_K$ if and only if it is a parametric entropy solution for $\mathcal{E}_\mathcal{C}$.*

Proof. For the proof, see Lemma 4.1 in [14] and the discussion which follows. □

Theorem 2.7. *If, for ρ-almost all $\boldsymbol{\xi}$, the initial data $u_0(\cdot, \boldsymbol{\xi}) \in \mathcal{L}^\infty(\mathbb{R}^n)$ and if $\mathbf{f}$ satisfies Assumption 1.1, then problem* (1) *has a unique parametric entropy solution u for $\mathcal{E}_K$, or equivalently, $\mathcal{E}_\mathcal{C}$. Moreover it satisfies for all $t \in \mathbb{R}_+$ and for ρ-almost all $\boldsymbol{\xi}$,*

$$\|u(t, \cdot, \boldsymbol{\xi})\|_{\mathcal{L}^\infty(\mathbb{R}^n)} \leq \|u_0(\cdot, \boldsymbol{\xi})\|_{\mathcal{L}^\infty(\mathbb{R}^n)}. \tag{3}$$

Proof. From [14, Theorem 5.2], we have that, for ρ-almost all $\boldsymbol{\xi}$, there exists a unique solution $u(\cdot, \cdot, \boldsymbol{\xi}) \in \mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$ for $\mathcal{E}_\mathcal{C}$. Then, from Proposition 2.6, we have that, for ρ-almost all $\boldsymbol{\xi}$, there exists a unique solution $u(\cdot, \cdot, \boldsymbol{\xi}) \in \mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$ for $\mathcal{E}_K$. The stability property (3) is deduced from [8, Theorem 6.2.7]. □

Remark 2.8. Note that Theorem 2.7 does not provide any information on the measurability of u. Under the assumption that $\Xi \ni \boldsymbol{\xi} \mapsto u(\cdot,\cdot,\boldsymbol{\xi}) \in \mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$ is Bochner measurable and that $u_0 \in \mathcal{L}^\infty(\Xi;\mathcal{L}^\infty(\mathbb{R}^n))$, Theorem 2.7 allows to deduce

$$\|u\|_{\mathcal{L}^\infty(\mathbb{R}_+\times\mathbb{R}^n\times\Xi)} \leq \|u_0\|_{\mathcal{L}^\infty(\mathbb{R}^n\times\Xi)}.$$

We refer to [31, Theorem 3.3] that provides measurability properties of u under additional assumptions on u_0.

Remark 2.9. We might hope that imposing that $u_0 \in \mathcal{L}^\infty(\Xi;\mathcal{L}^\infty(\mathbb{R}^n))$ may be sufficient to have that $\boldsymbol{\xi} \mapsto u(\cdot,\cdot,\boldsymbol{\xi})$ is Bochner measurable, but this has not been proved yet.

2.2 Weak-parametric entropy solutions

The next notion of solution is weaker. While the parametric entropy solution adopts a pointwise point of view in the parameter domain, the following notion of solution is deduced by integration over the parameter domain.

Definition 2.10 (Weak-parametric entropy solution)**.** Consider a family of entropy pairs $\mathcal{E}$. Let $u_0 \in \mathcal{L}^\infty(\mathbb{R}^n\times\Xi)$ and $\mathbf{f}$ satisfy Assumption 1.1. A measurable function $u : \mathbb{R}_+ \times \mathbb{R}^n \times \Xi \to \mathbb{R}$ in $\mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n \times \Xi)$ is called a weak-parametric entropy solution for $\mathcal{E}$ if, for all $(\eta,\mathbf{q}) \in \mathcal{E}$ and all non-negative test functions $\phi \in \mathcal{C}(\Xi;\mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n))$, it satisfies

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}^n}\int_{\Xi} \left(\partial_t\phi\eta(u) + \nabla_{\mathbf{x}}\phi\cdot\mathbf{q}(u,\boldsymbol{\xi})\right) d\rho(\boldsymbol{\xi})d\mathbf{x}dt + \int_{\mathbb{R}^n}\int_{\Xi} \phi\eta(u_0)d\rho(\boldsymbol{\xi})d\mathbf{x} \geq 0. \tag{4}$$

It is at first glance a weaker notion of solution, but we shall see that under certain assumptions, both notions of parametric entropy solution and weak-parametric entropy solution coincide.

Theorem 2.11. *Assume that $u_0 \in \mathcal{L}^\infty(\Xi,\mathcal{L}^\infty(\mathbb{R}^n))$ and $\mathbf{f}$ satisfies Assumption 1.1. A function u, such that $\boldsymbol{\xi} \mapsto u(\cdot,\cdot,\boldsymbol{\xi})$ is Bochner measurable, is a parametric entropy solution for $\mathcal{E}_K$ if and only if it is a weak-parametric entropy solution for $\mathcal{E}_K$.*

Proof. Let u be a parametric entropy solution for $\mathcal{E}_K$, and let $v \in \mathbb{R}$ and $\phi \in \mathcal{C}(\Xi;\mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n))$. From Remark 2.8, and since $u_0 \in \mathcal{L}^\infty(\Xi,\mathcal{L}^\infty(\mathbb{R}^n))$ and $\boldsymbol{\xi} \mapsto u(\cdot,\cdot,\boldsymbol{\xi})$ is Bochner measurable, we have that $u \in \mathcal{L}^\infty(\mathbb{R}_+\times\mathbb{R}^n\times\Xi)$. For ρ-almost all $\boldsymbol{\xi}$, $\phi(\cdot,\cdot,\boldsymbol{\xi}) \in \mathcal{C}_c^1(\mathbb{R}_+\times\mathbb{R}^n)$, thus u verifies equation (2) for $\psi = \phi(\cdot,\cdot,\boldsymbol{\xi})$. Let us integrate equation (2) on Ξ. First, let us consider the terms where η_v appears. From Remark 2.8, u is essentially bounded. Since η_v is continuous, $\eta_v \circ u$ is also essentially bounded. Since ϕ and its derivative are continuous in $\boldsymbol{\xi}$ and Ξ is a compact set, the terms where η_v appears are integrable in $\boldsymbol{\xi}$. Recalling the definition of $\mathbf{q}_v$ in Definition 2.3, we have that, for all $y \in \mathbb{R}$, ρ-almost everywhere, $\|\mathbf{q}_v(y,\boldsymbol{\xi})\| \leq \|\mathbf{f}(y,\boldsymbol{\xi})\| + \|\mathbf{f}(v,\boldsymbol{\xi})\|$. From Assumption 1.1, and since from the same argument as for the terms where η_v appears, u is essentially bounded, there exists $C, C_v \in \mathbb{R}$ such that $\|\mathbf{q}_v(u,\boldsymbol{\xi})\| \leq C + C_v$, ρ-almost everywhere. Thus, $\boldsymbol{\xi} \mapsto \int_{\mathbb{R}_+}\int_{\mathbb{R}^n} \nabla_{\mathbf{x}}\phi\cdot\mathbf{q}(u,\boldsymbol{\xi})d\mathbf{x}dt$ is integrable and integrating on Ξ yields equation (4).

Conversely, let $u \in \mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R} \times \boldsymbol{\Xi})$ be a weak-parametric entropy solution, $v \in \mathbb{R}$ and $\phi \in \mathcal{C}(\boldsymbol{\Xi}; \mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n))$ such that $\phi(t, \mathbf{x}, \boldsymbol{\xi}) := \psi(t, \mathbf{x})\gamma(\boldsymbol{\xi})$ where $\psi \in \mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n)$ and $\gamma \in \mathcal{C}(\boldsymbol{\Xi})$. The function u then verifies (4) for our particular choice of ϕ. Inequality (4) can be rewritten as $\int_{\boldsymbol{\Xi}} I(\boldsymbol{\xi})\gamma(\boldsymbol{\xi}) d\rho(\boldsymbol{\xi}) \geq 0$. Since I is ρ-measurable, $\mathbb{1}_{I<0}$ is ρ-measurable. Moreover, from [2, Theorem 12.7], since ρ is a finite Borel measure and $\boldsymbol{\Xi}$ is a Polish space, we have that ρ is a regular measure. Thus, there exists a sequence $\gamma_n \in \mathcal{C}(\boldsymbol{\Xi})$ such that $\|\mathbb{1}_{I<0} - \gamma_n\|_{\mathcal{L}^1(\boldsymbol{\Xi})} \to 0$ as $n \to \infty$ and $\int_{\boldsymbol{\Xi}} I(\boldsymbol{\xi}) \mathbb{1}_{I<0}(\boldsymbol{\xi}) d\rho(\boldsymbol{\xi}) \geq 0$. Yet, $I \mathbb{1}_{I<0} \leq 0$ ρ-almost everywhere. Thus, $I \geq 0$ ρ-almost everywhere and it gives us that u is a parametric entropy solution, which concludes the proof.

□

2.3 Measure-valued solutions

Following DiPerna [10], previous notions of solutions are extended to the weaker case of measure-valued solutions thanks to the notion of Young measure.

Definition 2.12 (Young measure)**.** A Young measure on a Euclidean space $\mathcal{X}$ is a map $\mu : \mathcal{X} \to \mathcal{P}(\mathbb{R}), \tau \mapsto \mu_\tau$ such that for all $g \in \mathcal{C}_0(\mathbb{R})$ the function $\tau \mapsto \int_{\mathbb{R}} g(y)\mu_\tau(dy)$ is measurable.

From this, we can seek an even weaker notion of solution that is a Young measure $\mu_{(t,\mathbf{x},\boldsymbol{\xi})}$ which satisfies the following Cauchy problem:

$$\partial_t \langle \mu_{(t,\mathbf{x},\boldsymbol{\xi})}, \mathrm{id}_{\mathbb{R}} \rangle + \mathrm{div}_{\mathbf{x}} \langle \mu_{(t,\mathbf{x},\boldsymbol{\xi})}, \mathbf{f}(\cdot, \boldsymbol{\xi}) \rangle = 0, \quad (t, \mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}_+ \times \mathbb{R}^n \times \boldsymbol{\Xi}, \tag{5a}$$

$$\mu_{(0,\mathbf{x},\boldsymbol{\xi})} = \sigma_0, \quad (\mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}^n \times \boldsymbol{\Xi}, \tag{5b}$$

where $\langle \cdot, \cdot \rangle$ denotes the integration of a (vector-valued) function $g \in \mathcal{C}(\mathbb{R}; \mathbb{R}^k)$ against a measure $\mu \in \mathcal{M}(\mathbb{R})$, defined by

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

while $\sigma_0 = \delta_{u_0}$. Equation (5a) has to be understood in a weak entropy sense, as explained in the following.

Definition 2.13 (Parametric entropy measure-valued (MV) solution)**.** Consider a family of entropy pairs $\mathcal{E}$. Let σ_0 be Young measure on $\mathbb{R}^n \times \boldsymbol{\Xi}$, and let $\mathbf{f}$ satisfying Assumption 1.1. A Young measure μ is a parametric entropy MV solution to (5) for $\mathcal{E}$, if, for a family of entropy pairs $\mathcal{E}$, for all $(\eta, \mathbf{q}) \in \mathcal{E}$ and all non-negative test functions $\phi \in \mathcal{C}(\boldsymbol{\Xi}; \mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n))$, it satisfies

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}^n} \int_{\boldsymbol{\Xi}} \left(\partial_t \phi(t, \mathbf{x}, \boldsymbol{\xi}) \langle \mu_{(t,\mathbf{x},\boldsymbol{\xi})}, \eta_v \rangle + \nabla_{\mathbf{x}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \cdot \langle \mu_{(t,\mathbf{x},\boldsymbol{\xi})}, \mathbf{q}_v \rangle \right) d\rho(\boldsymbol{\xi}) d\mathbf{x} dt$$
$$+ \int_{\mathbb{R}^n} \int_{\boldsymbol{\Xi}} \phi(0, \mathbf{x}, \boldsymbol{\xi}) \langle \sigma_0, \eta_v \rangle d\rho(\boldsymbol{\xi}) d\mathbf{x} \geq 0. \tag{6}$$

With the injection

$$\mathcal{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n \times \boldsymbol{\Xi}) \to (\mathbb{R}_+ \times \mathbb{R}^n \times \boldsymbol{\Xi} \to \mathcal{P}(\mathbb{R}))$$
$$u \mapsto \left((t, \mathbf{x}, \boldsymbol{\xi}) \mapsto \delta_{u(t,\mathbf{x},\boldsymbol{\xi})} \right),$$

we notice that, under the condition that $\sigma_0 = \delta_{u_0}$, weak-parametric entropy solutions are parametric entropy MV solutions, but without further assumptions, parametric entropy MV solutions are not necessarily weak-parametric entropy solutions. However, the following result shows that the parametric entropy MV solution can be concentrated on the graph of the weak-parametric entropy solution.

Theorem 2.14. *Let $u_0 \in \mathcal{L}^\infty(\mathbb{R}^n \times \Xi)$ and $\mathbf{f}$ satisfy Assumption 1.1. Let u be the unique weak parametric entropy solution for $\mathcal{E}_K$ and μ be a parametric entropy MV solution for $\mathcal{E}_K$. If ρ-almost everywhere $\sigma_0 = \delta_{u_0(\cdot)}$, then ρ-almost everywhere $\mu = \delta_{u(\cdot)}$.*

Proof. First, we may note that from the same arguments that those presented in the second part of the proof of Theorem 2.11, a parametric entropy MV solution μ for $\mathcal{E}_K$ verifies the following inequality, for all $v \in \mathbb{R}$, all non-negative test functions $\psi \in \mathcal{C}_c^1(\mathbb{R}_+ \times \mathbb{R}^n)$ and ρ-almost all $\boldsymbol{\xi}$:

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}^n} \left(\partial_t \psi(t,\mathbf{x})\langle \mu_{(t,\mathbf{x},\boldsymbol{\xi})}, \eta_v\rangle + \nabla_{\mathbf{x}}\psi(t,\mathbf{x}) \cdot \langle \mu_{(t,\mathbf{x},\boldsymbol{\xi})}, \mathbf{q}_v\rangle\right) d\mathbf{x}dt + \int_{\mathbb{R}^n} \psi(0,\mathbf{x})\langle \sigma_0, \eta_v\rangle d\mathbf{x} \geq 0. \quad (7)$$

Then, for ρ-almost all $\boldsymbol{\xi}$, $\mu(\cdot,\cdot,\boldsymbol{\xi})$ is an entropy MV solution of the initial problem with a fixed parameter $\boldsymbol{\xi}$, that is a parameter-independent problem studied in [29]. Then, from [29, Theorem 1] and [10], we have that ρ-almost everywhere, if $(\sigma_0)_{(\cdot,\boldsymbol{\xi})} = \delta_{u(\cdot,\boldsymbol{\xi})}$, then $\mu_{(\cdot,\cdot,\boldsymbol{\xi})} = \delta_{u(\cdot,\cdot,\boldsymbol{\xi})}$, with $u(\cdot,\cdot,\boldsymbol{\xi})$ the weak parametric entropy solution. □

2.4 Restrictions to compact hypercubes

In order to extend the strategy developed in [29], it is mandatory to work on compact sets. Whereas introducing compact domains in time and in the parameter set is trivial, the restriction to bounded space domains has to be carefully done. To simplify the setting and to avoid the problem of introducing boundary conditions to conservation laws, see for instance [4, 33, 32], we assume that the solution has no interaction with its boundary, i.e. the solution is known on the boundary of the spatial domain at any time. Let

$$\mathbf{T} := [0,T], \quad \mathbf{X} := [L_1,R_1] \times \cdots \times [L_n,R_n], \quad \Xi := [0,1]^p \quad (8)$$

be the respective domains of time t, space variable $\mathbf{x}$ and parameter $\boldsymbol{\xi}$ for fixed (but arbitrary) constants T, $(L_i)_{i=1}^n$ and $(R_i)_{i=1}^n$. The absence of interaction with the boundary is translated as follows: initial data u_0 that we consider are the restrictions to $\mathbf{X} \times \Xi$ of initial data defined on $\mathbb{R}^n \times \Xi$ such that, considering the associated weak parametric entropy solution u, there exists $\epsilon > 0$ such that in $\partial \boldsymbol{X}_\epsilon := (\partial\mathbf{X} + B(0,\epsilon)) \cap \mathbf{X}$, and for all $t \in \mathbf{T}$ and ρ-almost all $\boldsymbol{\xi}$, $u(t,\cdot,\boldsymbol{\xi}) = u_0(\cdot,\boldsymbol{\xi})$, i.e. the weak parametric entropy solution is stationary on $\partial \boldsymbol{X}_\epsilon$. This framework is the one we shall use in the following.

From (3), we can consider that u takes values in the following compact set

$$\mathbf{U} := [\underline{u}, \overline{u}], \quad (9)$$

where the bounds are $\underline{u} := \operatorname{ess\,inf}_{\mathbf{X},\boldsymbol{\Xi}} u_0$ and $\overline{u} := \operatorname{ess\,sup}_{\mathbf{X},\boldsymbol{\Xi}} u_0$.

This leads us to reformulate the problem on the restricted domain.

Proposition 2.15 (Parametric entropy measure-valued solution on compact hypercubes)**.** *Consider a family of entropy pair $\mathcal{E}$. Let $\mu : (t, \mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}_+ \times \mathbb{R}^n \times \boldsymbol{\Xi} \mapsto \mu_{(t,\mathbf{x},\boldsymbol{\xi})} \in \mathcal{M}_+(\mathbf{U})$ be a parametric entropy measure-valued solution for $\mathcal{E}$. Then it satisfies for all $(\eta, \mathbf{q}) \in \mathcal{E}$ and for all non-negative test functions $\phi \in \mathcal{C}(\boldsymbol{\Xi}; \mathcal{C}^1(\mathbf{T} \times \mathbf{X}))$,*

$$\begin{aligned}\int_{\mathbf{T}}\int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \left(\partial_t\phi(t,\mathbf{x},\boldsymbol{\xi})\langle\mu_{(t,\mathbf{x},\boldsymbol{\xi})},\eta\rangle + \nabla_{\mathbf{x}}\phi(t,\mathbf{x},\boldsymbol{\xi})\cdot\langle\mu_{(t,\mathbf{x},\boldsymbol{\xi})},\mathbf{q}\rangle\right) d\rho(\boldsymbol{\xi})d\mathbf{x}dt \\ + \int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(0,\mathbf{x},\boldsymbol{\xi})\langle\sigma_0,\eta\rangle d\rho(\boldsymbol{\xi})d\mathbf{x} - \int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(T,\mathbf{x},\boldsymbol{\xi})\langle\sigma_T,\eta\rangle d\rho(\boldsymbol{\xi})d\mathbf{x} \\ - \int_{\mathbf{T}}\int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(t,\boldsymbol{\partial}\mathbf{x},\boldsymbol{\xi})\langle\boldsymbol{\gamma},\mathbf{q}(\cdot,\boldsymbol{\xi})\rangle d\rho(\boldsymbol{\xi})d\mathbf{x}dt \geq 0 \quad (10)\end{aligned}$$

where σ_0 and σ_T are Young measures supported on $\mathbf{X} \times \boldsymbol{\Xi}$, and where $\boldsymbol{\gamma}$ is such that

$$\begin{aligned}&\int_{\mathbf{T}}\int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(t,\boldsymbol{\partial}\mathbf{x},\boldsymbol{\xi})\langle\boldsymbol{\gamma},\mathbf{q}(\cdot,\boldsymbol{\xi})\rangle d\rho(\boldsymbol{\xi})d(\mathbf{x})dt = \\ &\sum_{i=1}^{n}\int_{\mathbf{T}}\int_{\Gamma_{R,i}}\int_{\boldsymbol{\Xi}} \phi(t,\mathbf{x}_{L,i},\boldsymbol{\xi})\langle\gamma_{R,i},q_i(\cdot,\boldsymbol{\xi})\rangle d\rho(\boldsymbol{\xi})d\mathbf{x}dt \\ &-\sum_{i=1}^{n}\int_{\mathbf{T}}\int_{\Gamma_{R,i}}\int_{\boldsymbol{\Xi}} \phi(t,\mathbf{x}_{R,i},\boldsymbol{\xi})\langle\gamma_{L,i},q_i(\cdot,\boldsymbol{\xi})\rangle d\rho(\boldsymbol{\xi})d\mathbf{x}dt,\end{aligned}$$

where for each $1 \leq i \leq n$, $\gamma_{L,i}$ and $\gamma_{R,i}$ are boundary measures supported on $\mathbf{T} \times \Gamma_{L,i} \times \boldsymbol{\Xi}$ and $\mathbf{T} \times \Gamma_{R,i} \times \boldsymbol{\Xi}$ respectively, with $\Gamma_{L,i} = \{\mathbf{x} \in \partial\mathbf{X} : x_i = L_i\}$ and $\Gamma_{R,i} = \{\mathbf{x} \in \partial\mathbf{X} : x_i = R_i\}$, $\mathbf{x}_{B,i}$ denotes the vector $(x_1, \ldots, x_{i-1}, \Gamma_{B,i}, x_{i+1}, \ldots, x_n)$ for $B \in \{L, R\}$.

Lemma 2.16. *Consider a family of entropy pairs $\mathcal{E}$ such that either $(\mathrm{id}, \mathbf{f}) \in \mathcal{E}$ or $\mathcal{E} = \mathcal{E}_K$. Let μ be a parametric entropy MV solution for $\mathcal{E}$. Then for all test functions $\phi \in \mathcal{C}(\boldsymbol{\Xi}; \mathcal{C}^1(\mathbf{T} \times \mathbf{X}))$, it satisfies*

$$\begin{aligned}\int_{\mathbf{T}}\int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \left(\partial_t\phi(t,\mathbf{x},\boldsymbol{\xi})\langle\mu_{(t,\mathbf{x},\boldsymbol{\xi})},\mathrm{id}\rangle + \nabla_{\mathbf{x}}\phi(t,\mathbf{x},\boldsymbol{\xi})\langle\mu_{(t,\mathbf{x},\boldsymbol{\xi})},\mathbf{f}\rangle\right) d\rho(\boldsymbol{\xi})d\mathbf{x}dt \\ + \int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(0,\mathbf{x},\boldsymbol{\xi})\langle\sigma_0,\mathrm{id}\rangle d\rho(\boldsymbol{\xi})d\mathbf{x} - \int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(T,\mathbf{x},\boldsymbol{\xi})\langle\sigma_T,\mathrm{id}\rangle d\rho(\boldsymbol{\xi})d\mathbf{x} \\ - \int_{\mathbf{T}}\int_{\mathbf{X}}\int_{\boldsymbol{\Xi}} \phi(t,\mathbf{x},\boldsymbol{\xi})\langle\boldsymbol{\gamma},\mathbf{f}(\cdot,\boldsymbol{\xi})\rangle d\rho(\boldsymbol{\xi})d\mathbf{x}dt = 0. \quad (11)\end{aligned}$$

Proof. The proof of this lemma is discussed in [12], and the case of the Kruzhkov's entropies is retrieved thanks to the boundedness of $\mathbf{U}$. □

Remark 2.17. Our numerical method shall not use all inequalities (10). Therefore, imposing (11) as an additional constraint may be beneficial in practice, see further discussion in Section 3.

Remark 2.18 (Imposing constraints on the boundary)**.** To ensure concentration of $\mu_{(t,x,\boldsymbol{\xi})}$, in addition to the condition $\sigma_0 = \delta_{u_0(\cdot)}$, one may impose conditions on the boundary measures $(\gamma_{L,i})_{i=1}^n$ and $(\gamma_{R,i})_{i=1}^n$. The choice of boundary condition allows to ensure the absence of interaction with the boundary. We shall make the assumption that the trace of u_0 on $\Gamma_{B,i}$, noted $\gamma_{B,i}(u_0)$ exists for all $B \in \{L, R\}$ and all $1 \leq i \leq n$, and we at the same time notice that this trace does not depend on $\boldsymbol{\xi} \in \boldsymbol{\Xi}$. We then want to impose that $\gamma_{B,i}(\mathbf{x}) = \delta_{\gamma_{B,i}(u_0)(\mathbf{x})}$ for almost all $\mathbf{x} \in \mathbf{X}$, for all $1 \leq i \leq n$ and for $B \in \{L, R\}$.

Let $\nu \in \mathcal{M}(\mathbf{K})_+$ with $\mathbf{K} := \mathbf{T} \times \mathbf{X} \times \boldsymbol{\Xi} \times \mathbf{U}$ defined by

$$d\nu(t, \mathbf{x}, \boldsymbol{\xi}, y) = dtd\mathbf{x}d\rho(\boldsymbol{\xi})\mu_{(t,\mathbf{x},\boldsymbol{\xi})}(dy) \tag{12}$$

where μ is a parametric entropy MV solution. The measure ν, called occupation measure (see [26]), has $\lambda_{\mathbf{T}} \otimes \lambda_{\mathbf{X}} \otimes \rho$ for marginal in $(t, \mathbf{x}, \boldsymbol{\xi})$, and $\mu_{(t,\mathbf{x},\boldsymbol{\xi})}$ as the conditional measure in y given $(t, \mathbf{x}, \boldsymbol{\xi})$. In the case where $\mu_{(t,\mathbf{x},\boldsymbol{\xi})}(dy) = \delta_{u(t,\mathbf{x},\boldsymbol{\xi})}(dy)$, ν is supported on the graph of the function u. We also introduce the time boundary measures

$$d\nu_0(t, \mathbf{x}, \boldsymbol{\xi}, y) := \delta_0(dt)d\mathbf{x}d\rho(\boldsymbol{\xi})\sigma_0(dy), \quad d\nu_T(t, \mathbf{x}, \boldsymbol{\xi}, y) := \delta_T(dt)d\mathbf{x}d\rho(\boldsymbol{\xi})\sigma_T(dy) \tag{13}$$

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

$$d\nu_{L,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) := dt\delta_{\Gamma_{L,i}}(d\mathbf{x})d\rho(\boldsymbol{\xi})\gamma_{L,i}(dy), \tag{14}$$

$$d\nu_{R,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) := dt\delta_{\Gamma_{R,i}}(d\mathbf{x})d\rho(\boldsymbol{\xi})\gamma_{R,i}(dy) \tag{15}$$

whose supports are given by $\mathbf{K}_{L,i} := \mathbf{T} \times \Gamma_{L,i} \times \boldsymbol{\Xi} \times \mathbf{U}$ and $\mathbf{K}_{R,i} := \mathbf{T} \times \Gamma_{R,i} \times \boldsymbol{\Xi} \times \mathbf{U}$ respectively, for $1 \leq i \leq n$. For conciseness, we shall define the collection of measures $\boldsymbol{\nu} := \{\nu, \nu_0, \nu_T, (\nu_{L,i})_{i=1}^n, (\nu_{R,i})_{i=1}^n\}$.

It is known that measures with compact support are fully characterized by their moments, see e.g. [24, p.52]. Thus, marginal constraints on occupation measures $\boldsymbol{\nu}$ will be imposed through their moments, see details in Appendix A.

The introduction of the measure ν allows us to rewrite the constraints (11) and (10). The equation (11) can be put in the following form

$$\begin{aligned} F(\phi, \boldsymbol{\nu}) := &\int_{\mathbf{K}} \left(\partial_t \phi(t, \mathbf{x}, \boldsymbol{\xi}) y + \nabla_{\mathbf{x}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \cdot \mathbf{f}(y)\right) d\nu(t, \mathbf{x}, \boldsymbol{\xi}, y) \\ &+ \int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) y d\nu_0(t, \mathbf{x}, \boldsymbol{\xi}, y) - \int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) y d\nu_T(t, \mathbf{x}, \boldsymbol{\xi}, y) \\ + \sum_{i=1}^n &\left(\int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) f_i(y) d\nu_{L,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) - \int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) f_i(y) d\nu_{R,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) \right) = 0, \end{aligned} \tag{16}$$

where $\phi \in \mathcal{C}(\boldsymbol{\Xi}; \mathcal{C}^1(\mathbf{T} \times \mathbf{X}))$, and the equation (10) can be written

$$\begin{aligned} G(\phi, \boldsymbol{\nu}, \eta, \boldsymbol{q}) := & \int_{\mathbf{K}} \left(\partial_t \phi(t, \mathbf{x}, \boldsymbol{\xi}) \eta(y) + \nabla_{\mathbf{x}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \cdot \mathbf{q}(y)\right) d\nu(t, \mathbf{x}, \boldsymbol{\xi}, y) \\ & + \int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \eta(y) d\nu_0(t, \mathbf{x}, \boldsymbol{\xi}, y) - \int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \eta(y) d\nu_T(t, \mathbf{x}, \boldsymbol{\xi}, y) \\ + \sum_{i=1}^{n} & \left(\int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) q_i(y) d\nu_{L,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) - \int_{\mathbf{K}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) q_i(y) d\nu_{R,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) \right) \geq 0, \end{aligned} \tag{17}$$

for entropy pairs $(\eta, \mathbf{q})$ in a family $\mathcal{E}$ and $\phi \in \mathcal{C}(\boldsymbol{\Xi}; \mathcal{C}^1(\mathbf{T} \times \mathbf{X}))$ non-negative test functions.

3 Moment-SOS method for measure-valued solutions on compact sets

In the previous section, we introduced parametric measure-valued (MV) solutions for scalar hyperbolic equations, that are defined by equations (16)-(17). The aim of this section is to express these equations as constraints on the moments of the occupation measure and to explain how to approximate these moments based on the moment-SOS (Lasserre's) hierarchy [24]. For that, we require the assumption that $\mathbf{f} : \mathbb{R} \times \boldsymbol{\Xi}$ is a polynomial function.

We will see in the next section how to extract from these moments some information on the solution u of the initial problem.

3.1 From weak formulations to moment constraints

The following lemma, derived from [29, Lemma 1], relies on density arguments, together with the fact that we are working with compact sets.

Lemma 3.1. *Let $\{\phi^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathbb{N}^{n+p+1}}$ be a polynomial basis on $\mathbf{T} \times \mathbf{X} \times \boldsymbol{\Xi}$. Then equation (16) is equivalent to*

$$F(\phi^{\boldsymbol{\alpha}}, \boldsymbol{\nu}) = 0 \tag{18}$$

for all $\boldsymbol{\alpha} \in \mathbb{N}^{n+p+1}$, where F is defined in (16).

For $\mathbf{f}$ a polynomial function, (18) provides constraints on linear combinations of moments of measures $\boldsymbol{\nu}$. In the case of a family of polynomial entropy pairs $\mathcal{E} \subseteq \mathcal{E}_P$, we can also express (17) as constraints on the moments of measures $\boldsymbol{\nu}$.

Lemma 3.2. *Assume $\{\phi^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{F}}$ for $\mathcal{F} \subset \mathbb{N}^{n+p+1}$ is a countable family of polynomials on $\mathbf{T} \times \mathbf{X} \times \boldsymbol{\Xi}$ such that any non-negative polynomial can be decomposed on this family, with positive coefficients. Then, equation (17) is equivalent to*

$$G(\phi^{\boldsymbol{\alpha}}, \boldsymbol{\nu}, \eta, \mathbf{q}) \geq 0 \tag{19}$$

for all $\boldsymbol{\alpha} \in \mathcal{F}$ and all $(\eta, \mathbf{q}) \in \mathcal{E}$.

Remark 3.3. Since, as stated in Lemma 2.16, equation (17) implies equation (16), then equation (19) implies equation (18) with an appropriate family of entropy pairs. It thus may seem redundant to enforce both, but, in the approximation method, the family of polynomials in Lemma 3.2 will be reduced, so that this implication is no more guaranteed and imposing (18) as additional constraints may be beneficial.

The case $\mathcal{E} = \mathcal{E}_K$ ensures concentration of the measure, as seen in Theorem 2.14, but we are faced with two issues: first, taking into account an uncountable family of functions parametrised by $v \in \mathbf{U}$ 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 introduce as new unknowns Borel measures ϑ^+ and ϑ^-, whose supports are respectively defined by

$$\operatorname{supp}(\vartheta^+) = \mathbf{K}^+ := \{(t, \mathbf{x}, \boldsymbol{\xi}, y, v) \in \mathbf{K} \times \mathbf{U} : y \geq v\},$$
$$\operatorname{supp}(\vartheta^-) = \mathbf{K}^- := \{(t, \mathbf{x}, \boldsymbol{\xi}, y, v) \in \mathbf{K} \times \mathbf{U} : y \leq v\},$$

and impose the condition that $\nu \otimes \lambda_{\mathbf{U}} = \vartheta^+ + \vartheta^-$, which can be expressed as constraints between moments of ν, ϑ^+ and ϑ^-. Similarly, we introduce time boundary measures ϑ_0^+, ϑ_0^-, ϑ_T^+ and ϑ_T^-, space boundary measures $(\vartheta_{L,i}^+)_{i=1}^n$, $(\vartheta_{L,i}^-)_{i=1}^n$, $(\vartheta_{R,i}^+)_{i=1}^n$ and $(\vartheta_{R,i}^-)_{i=1}^n$, and the corresponding constraints with measures $\boldsymbol{\nu}$. All those definitions are plainly written in Appendix B. We shall once again introduce a collection of measures

$$\boldsymbol{\vartheta} := (\vartheta^+, \vartheta^-, \vartheta_0^+, \vartheta_0^-, \vartheta_T^+, \vartheta_T^-, (\vartheta_{L,i}^+)_{i=1}^n, (\vartheta_{L,i}^-)_{i=1}^n, (\vartheta_{R,i}^+)_{i=1}^n, (\vartheta_{R,i}^-)_{i=1}^n).$$

From [29, Lemma 2], equation (17) is equivalent to

$$\begin{aligned}
H(\phi, \boldsymbol{\vartheta}) := &\int_{\mathbf{K}} \theta(v) \left(\partial_t \phi(t, \mathbf{x}, \boldsymbol{\xi})(y - v) + \nabla_{\mathbf{x}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \cdot (\mathbf{f}(y) - \mathbf{f}(v))\right) d\vartheta^+ \\
&+ \int_{\mathbf{K}} \theta(v) \left(\partial_t \phi(t, \mathbf{x}, \boldsymbol{\xi})(v - y) + \nabla_{\mathbf{x}} \phi(t, \mathbf{x}, \boldsymbol{\xi}) \cdot (\mathbf{f}(v) - \mathbf{f}(y))\right) d\vartheta^- \\
&+ \int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(y - v) d\vartheta_0^+ + \int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(v - y) d\vartheta_0^- \\
&- \int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(y - v) d\vartheta_T^+ - \int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(v - y) d\vartheta_T^- \\
+ \sum_{i=1}^n \Bigg(&\int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(f_i(y) - f_i(v)) d\vartheta_L^+ + \int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(f_i(v) - f_i(y)) d\vartheta_L^- \\
- &\int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(f_i(y) - f_i(v)) d\vartheta_R^+ - \int_{\mathbf{K}} \theta(v) \phi(t, \mathbf{x}, \boldsymbol{\xi})(f_i(v) - f_i(y)) d\vartheta_R^- \Bigg) \geq 0,
\end{aligned} \tag{20}$$

for all non-negative test functions $\phi \in \mathcal{C}^1(\mathbf{T} \times \mathbf{X}) \otimes \mathcal{C}(\boldsymbol{\Xi})$ and all non-negative test functions $\theta \in \mathcal{C}(\mathbf{U})$.

Lemma 3.4. *Assume $\{\phi^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{F}}$ is a countable family of polynomials on $\mathbf{T} \times \mathbf{X} \times \boldsymbol{\Xi} \times \mathbf{U}$ such that any non-negative polynomial can be decomposed on this family with positive coefficients. Then* (20) *is equivalent to*

$$H(\phi^{\boldsymbol{\alpha}}, \boldsymbol{\vartheta}) \geq 0 \tag{21}$$

for all $\boldsymbol{\alpha} \in \mathcal{F}$.

Proof. The proof relies on density arguments. □

A particular family $\{\phi^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^{2(n+p+2)}}$ satisfying the assumption that any non-negative polynomial can be decomposed on this family with positive coefficients is given by

$$\phi^{\boldsymbol{\alpha}}(t,\mathbf{x},\boldsymbol{\xi},v) := t^{\alpha_1}(T-t)^{\alpha_2}\prod_{i=1}^{n}((x_i-L_i)^{\alpha_{2i+1}}(R_i-x_i)^{\alpha_{2(i+1)}})$$
$$\prod_{i=1}^{p}((\xi_i)^{\alpha_{2i+2n+1}}(1-\xi_i)^{\alpha_{2(i+n+1)}})\,(v-\underline{u})^{\alpha_{2n+2p+3}}(\overline{u}-v)^{\alpha_{2(n+p+2)}}$$

for $\boldsymbol{\alpha} \in \mathbb{N}^{2(n+p+2)}$. The proof follows the one of the Lemma 3 in [29], and uses Handelman's Positivstellensatz [29].

3.2 Generalized Moment Problem

Roughly speaking, the Generalized Moment Problem (GMP) is an infinite-dimensional linear optimization problem on finitely many Borel measures $\nu_i \in \mathcal{M}(\mathbf{K}_i)_+$, with $\mathbf{K}_i \subseteq \mathbb{R}^{n_i}$, with $i = 1, ..., N$ 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 some criterion. In full generality, the GMP is intractable, but if all $\mathbf{K}_i$ are basic semi-algebraic sets[1] and the integrands are polynomials, 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 semi-definite programs[2] of increasing size. Convergence of this numerical scheme is guaranteed by invoking powerful results from real algebraic geometry, essentially positivity certificates, and further developed for many classical cases in [37, 20].

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 $i = 1, \ldots, N$ and countably many $k = 1, 2, \ldots$. The GMP is the following optimization problem over measures:

$$\begin{aligned}\inf_{\nu_1,\ldots,\nu_N} &\sum_{i=1}^{N}\int_{\mathbf{K}_i} h_i d\nu_i =: \rho^* \\ \text{s.t.} &\sum_{i=1}^{N}\int_{\mathbf{K}_i} h_{i,k} d\nu_i \le b_k, \quad k = 1, 2, \ldots \\ &\nu_i \in \mathcal{M}(\mathbf{K}_i)_+, \quad i = 1, \ldots, N.\end{aligned} \tag{22}$$

3.3 From measures to moments and their approximation

Instead of optimizing over the measures in problem (22), we optimize over their moments. For simplicity and clarity of exposition, we describe the approach in

[1] A basic semi-algebraic set is defined by $\{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) \ge 0, \forall i = 1, \ldots, m\}$ where $m \in \mathbb{N}$ and $f_1, \ldots, f_m$ are polynomials.

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

the case of a single unknown measure ν, but it easily extends to the case of several measures. Let us consider the simplified GMP

$$\begin{aligned} &\inf_{\nu} \int_{\mathbf{K}} h d\nu := \rho^* \\ &\text{s.t.} \int_{\mathbf{K}} h_k d\nu \le b_k, \quad k = 1, 2, \dots \\ &\qquad \nu \in \mathcal{M}(\mathbf{K})_+, \end{aligned} \tag{23}$$

where $\mathbf{K}$ is a compact set in $\mathbb{R}^n$, $h \in \mathbb{R}[\mathbf{w}]$, $h_k \in \mathbb{R}[\mathbf{w}]$ and $b_k \in \mathbb{R}$ for all $k = 1, 2, \dots$. The moment sequence $\mathbf{z} = (\mathbf{z}_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ of a measure $\nu \in \mathcal{M}(\mathbf{K})_+$ is defined by

$$\mathbf{z}_{\boldsymbol{\alpha}} = \int_{\mathbf{K}} \mathbf{w}^{\boldsymbol{\alpha}} d\nu, \quad \boldsymbol{\alpha} \in \mathbb{N}^n. \tag{24}$$

Similarly, given a sequence $\mathbf{z} = (\mathbf{z}_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$, if (24) holds for some $\nu \in \mathcal{M}(\mathbf{K})_+$ we say that the sequence has the representing measure ν. Recall that measures on compact sets are uniquely characterized by their moments (see [24, p. 52]).

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

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

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

$$\ell_{\mathbf{z}}(p) := \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} \mathbf{p}_{\boldsymbol{\alpha}} \mathbf{z}_{\boldsymbol{\alpha}} \tag{25}$$

and $\ell_{\mathbf{z}}$ is called the Riesz functional.

Theorem 3.5 (Riesz-Haviland [24, Theorem 3.1])**.** *Let $\mathbf{K} \subseteq \mathbb{R}^n$ be closed. A real sequence $\mathbf{z} \in \mathbb{R}^{\mathbb{N}^n}$ is the moment sequence of some measure $\nu \in \mathcal{M}(\mathbf{K})_+$, i.e. $\mathbf{z}$ satisfies* (24)*, if and only if $\ell_{\mathbf{z}}(p) \ge 0$ for all $p \in \mathbb{R}[\mathbf{w}]$ non-negative on $\mathbf{K}$.*

Assuming that $\mathbf{K}$ is closed, we can reformulate thanks to this result the GMP (23) as a linear problem on moment sequences, namely

$$\begin{aligned} \inf_{\mathbf{z}} \quad & \ell_{\mathbf{z}}(h) = \rho^* \\ \text{s.t.} \quad & \ell_{\mathbf{z}}(h_k) \le b_k, \quad k = 1, 2, \dots \\ & \ell_{\mathbf{z}}(p) \ge 0, \text{ for all } p \in \mathbb{R}[\mathbf{w}] \text{ non-negative on } \mathbf{K}. \end{aligned} \tag{26}$$

Theorem 3.5 guarantees the equivalence between formulations (26) and (23). However, the latter reformulation is still numerically intractable.

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

$$\mathbf{K} = \{\mathbf{w} \in \mathbb{R}^n : g_1(\mathbf{w}) \ge 0, \dots, g_m(\mathbf{w}) \ge 0\} \tag{27}$$

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

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

Theorem 3.6 (Putinar's Positivstellensatz)**.** *If $p > 0$ on the basic semi-algebraic compact set $\mathbf{K}$ defined by (27) with $g_1(\mathbf{w}) := N - \sum_{i=1}^n w_i^2$, then $p = \sum_{j=0}^m s_j g_j$ for some SOS polynomials $s_j \in \mathbb{R}[\mathbf{w}], j = 0, 1, \ldots, m$.*

By a density argument, checking non-negativity of $\ell_{\mathbf{z}}$ on polynomials non-negative on $\mathbf{K}$ can be replaced by checking non-negativity only on polynomials that are strictly positive on $\mathbf{K}$ and hence on those that have a SOS representation as in Theorem 3.6.

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 $Q_d(g) \subset \mathbb{R}[\mathbf{w}]$ for $g = (g_0, \ldots, g_m)$ by

$$Q_d(g) := \left\{ \sum_{j=0}^m \sigma_j g_j : \deg(\sigma_j g_j) \leq 2d, \sigma_j \in \Sigma[\mathbf{w}]_d, j = 0, 1, \ldots, m \right\} \tag{28}$$

and observe that $Q_d(g)$ consists of polynomials which are non-negative on $\mathbf{K}$.

Let $\mathbf{b}_d(\mathbf{w}) := (\mathbf{w}^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leq d} \in \mathbb{R}[\mathbf{w}]^{n_d}$ be the vector of monomials of degree at most d. We recall that n_d denotes the binomial number $\binom{n+d}{n}$. For $j = 0, ..., m$, let $d_j = \lceil \deg(g_j)/2 \rceil$, 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 with polynomial entries $g_j \mathbf{b}_{d-d_j}(\mathbf{w}) \mathbf{b}_{d-d_j}^T(\mathbf{w})$. For $j = 0$ and $g_0 = 1$, the matrix $M_d(\mathbf{z}) = \ell_{\mathbf{z}}(\mathbf{b}_d \mathbf{b}_d^T)$ (where $\ell_{\mathbf{z}}$ is applied entrywise) is called the moment matrix. For any other value of j, it is called a localizing matrix. It turns out that, for all $j = 0, 1, \ldots, m$, $\ell_{\mathbf{z}}(g_j q^2) \geq 0$ for all $q \in \mathbb{R}[\mathbf{w}]_d$ if and only if $M_{d-d_j}(g_j \mathbf{z}) \succeq 0$, which are convex linear matrix inequalities in $\mathbf{z}$ and where $\succeq$ denotes the positive semi-definite (or Loewner) order.

Moment-SOS hierarchy. The following finite-dimensional semi-definite programming (SDP) problems are relaxations of the moment problem (26):

$$\begin{aligned} \inf_{\mathbf{z} \in \mathbb{R}^{n_{2d}}} \quad & \ell_{\mathbf{z}}(h) := \rho_d^* \\ \text{s.t.} \quad & \ell_{\mathbf{z}}(h_k) \leq b_k, \deg(h_k) \leq 2d, k = 1, 2, \ldots \\ & M_{d-d_j}(g_j \mathbf{z}) \succeq 0, j = 0, 1, \ldots, m \end{aligned} \tag{29}$$

and they are parametrized by the relaxation order $d \geq \max_{j=0,\ldots,m} d_j$.

Theorem 3.7 (Convergence of the moment-SOS hierarchy, [37, Theorem 5])**.** *Suppose that there exists $C > 0$ such that for any $d \in \mathbb{N}$, if $\mathbf{z}^d \in \mathbb{R}^{n_{2d}}$ is solution*

[3]This condition is slightly stronger than asking $\mathbf{K}$ to be a basic semi-algebraic compact set. However, the inequality $N - \sum_{i=1}^n w_i^2 \geq 0$ can always be added as a redundant constraint to the description of a basic semi-algebraic *compact* set. This condition has to be added because Putinar's result applies to a family of polynomials, and is not inherent to the set this family describes.

of (29)*, then* $z_{\mathbf{0}}^d \leq C$*, with* C *independent of* d*. Then there exists a sequence* $(\mathbf{z}^d)_d = ((\mathbf{z}^d_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n 2d})_d$ *such that* $\ell_{\mathbf{z}^d}(h) = \rho_d^*$ *and for all* $\boldsymbol{\alpha} \in \mathbb{N}^n$

$$\mathbf{z}^d_{\boldsymbol{\alpha}} \underset{d\to\infty}{\longrightarrow} \int_{\mathbf{K}} \mathbf{w}^{\boldsymbol{\alpha}} d\nu(\mathbf{w}). \tag{30}$$

In particular, one has $\rho_d^* \to \rho^*$ *as* $d \to \infty$*.*

3.4 Application to our problem

Entropy MV solution as a GMP. In the scalar hyperbolic case, the measures ν_i under consideration are from the collection $\boldsymbol{\nu}$, or $\boldsymbol{\nu}$ and $\boldsymbol{\vartheta}$ when considering Kruzhkov's entropies. The sets $\mathbf{K}_i$ all correspond to $\mathbf{K} = \mathbf{T}\times\mathbf{X}\times\boldsymbol{\Xi}\times\mathbf{U}$. The polynomials $h_{i,j}$ are given in (18) (conservation law), (19) when considering polynomial entropy pairs or (21) (and compatibility conditions between $\boldsymbol{\nu}$ and $\boldsymbol{\vartheta}$ (47) and similar equations) when considering Kruzhkov entropy pairs (entropy inequalities), and (38)-(41) (marginal constraints). For the sake of readibility, we shall only consider the case of polynomial entropies and a formulation only on measures $\boldsymbol{\nu}$.

We may also define an objective functional

$$\int_{\mathbf{K}} h d\nu + \int_{\mathbf{K}} h_0 d\nu_0 + \int_{\mathbf{K}} h_T d\nu_T + \sum_{i=1}^n \left(\int_{\mathbf{K}} h_{L,i} d\nu_{L,i} + \int_{\mathbf{K}} h_{R,i} d\nu_{R,i} \right), \tag{31}$$

with h, h_0, h_T, $(h_{L,i})_{i=1}^n$, $(h_{R,i})_{i=1}^n \in \mathbb{R}[t,\mathbf{x},\boldsymbol{\xi},y]$.

If the initial measure is concentrated on the graph of the initial condition and if, in addition, one imposes suitable boundary measures as exposed in Remark 2.18, then the choice of the objective functional is not crucial to recover the entropy MV solution of scalar hyperbolic PDE. Indeed, as a consequence of Theorem 2.14, the corresponding Young measure is concentrated: there is nothing to be optimized. However, our aim is to approximate the GMP by a finite dimensional optimization problem in order to solve it numerically and, then, the choice of the objective functional will impact the convergence of the corresponding relaxations. From experimental observations, two objective functionals seem to produce interesting results: the maximum of the opposite of the entropy constraints and the minimum of the trace of moment matrix. Choosing the latter seems to be a good heuristic: minimizing the nuclear norm of a matrix leads to reducing its rank (see [35]), which tends to favorise measures with localized support. However, there is up to date still no proof of a general effective functional.

Finally, one is able to define a GMP:

$$\begin{aligned} \inf_{\nu,\nu_T} \quad & (31) \text{ (objective functional)} \\ \text{s.t.} \quad & (18) \text{ (conservation law)}, \\ & (19) \text{ (entropy inequality)}, \\ & (38) - (41) \text{ (marginal constraints)}, \end{aligned} \tag{32}$$

where the infimum is taken over measures $\nu \in \mathcal{M}(\mathbf{K})_+, \nu_T \in \mathcal{M}(\mathbf{K}_T)_+$.

Remark 3.8. Note that the compact sets $\mathbf{T}$, $\mathbf{X}$, $\mathbf{\Xi}$ and $\mathbf{U}$ as defined before can be expressed as basic semi-algebraic compact sets:

$$\begin{aligned}
&\mathbf{T} = \{t \in \mathbb{R} : t(T-t) \geq 0\},\\
&\mathbf{X} = \{\mathbf{x} \in \mathbb{R} : (x_1 - L_1)(R_1 - x_1) \geq 0, \ldots, (x_n - L_n)(R_n - x_n) \geq 0\},\\
&\mathbf{\Xi} = \{\boldsymbol{\xi} \in \mathbf{\Xi} : \xi_1(1-\xi_1) \geq 0, \ldots, \xi_p(1-\xi_p) \geq 0\},\\
&\mathbf{U} = \{y \in \mathbb{R} : (y - \underline{u})(\overline{u} - y) \geq 0\}.
\end{aligned} \tag{33}$$

Theorem 3.7 extends to the case of multiple measures, as discussed in [24] and shown in [37]. Moreover, the constraint for $\boldsymbol{\alpha} = \mathbf{0}$ (see equation (38) in Appendix) yields the relaxed linear constraint $\mathbf{z_0} = \ell_{\mathbf{z}}(1) = \int_{\mathbf{T}\times\mathbf{X}\times\mathbf{\Xi}} dtdxd\rho(\boldsymbol{\xi}) \leq |\mathbf{T}||\mathbf{X}|$. Hence the hypothesis of Theorem 3.7 is verified.

Then, optimal solutions of the moment-SOS hierarchy (29) (adapted to the present context) converge to optimal solutions of (32) as d goes to infinity. In particular, one may extract the MV solution of (11), provided that σ_0, $\gamma_{L,i}$ and $\gamma_{R,j}$ are concentrated for $1 \leq i \leq n$, as already discussed in Remark 2.18.

3.5 Post-processing quantities of interest

We have seen in the previous section how to obtain approximate sequences $\mathbf{z}^d$ of moments of the measure ν on $\mathbf{K}$, such that $d\nu(t,\mathbf{x},\boldsymbol{\xi},y) = d\mu_{t,\mathbf{x},\boldsymbol{\xi}}(y)dtd\mathbf{x}d\rho(\boldsymbol{\xi})$ where μ is the measure-valued solution supported on the graph of the solution. In this section, we present how to construct an approximation of the function u thanks to the Christoffel-Darboux function and its ability to estimate the support of a measure (see [25] for further details). Also, we show how to obtain approximations of statistical moments of variables of interest that are functions of the solution, possibly using a moment completion technique and the Christoffel-Darboux function.

3.5.1 Approximation of the graph of the solution

We consider that we have obtained an approximation $\mathbf{z}^d$ of the moments of order $2d$ of the measure ν, which is a measure supported on the graph of the function $u(t,\mathbf{x},\boldsymbol{\xi})$. In order to approximate the function from the moments, we rely on an approximate Christoffel-Darboux function associated with the measure (that has to be carefully defined), which tends to take high values on the support of the measure. Thus, finding the minimizers of the approximate inverse Christoffel-Darboux function for given $(t,\mathbf{x},\boldsymbol{\xi}) \in \mathbf{T}\times\mathbf{X}\times\mathbf{\Xi}$ gives an approximation of $u(t,\mathbf{x},\boldsymbol{\xi})$. For $\mathbf{w} = (t,\mathbf{x},\boldsymbol{\xi},y) \in \mathbf{K}$, we let $\mathbf{b}_d(\mathbf{w})$ be a basis of monomials of order up to d and $M_d(\mathbf{z}^d) = \ell_{\mathbf{z}^d}(\mathbf{b}_d(\cdot)\mathbf{b}_d(\cdot)^T)$ be the corresponding moment matrix, that is the Gram matrix of the basis $\mathbf{b}_d(\mathbf{w})$ for the measure ν^d corresponding to $\mathbf{z}^d$. When $M_d(\mathbf{z}^d)$ is invertible, the inverse Christoffel-Darboux function is defined by

$$q_{\nu^d}(\mathbf{w}) = \mathbf{b}_d(\mathbf{w})^T M_d(\mathbf{z}^d)^{-1}\mathbf{b}_d(\mathbf{w}) = \sum_{i=1}^{n_d} \lambda_i^{-1}(\mathbf{b}_d(\mathbf{w})^T\mathbf{v}_i)^2$$

where the $(\lambda_i,\mathbf{v}_i)$ are eigenpairs of $M_d(\mathbf{z}^d)$, and the polynomials $p_i(\mathbf{w}) = \lambda_i^{-1/2}\mathbf{b}_d(\mathbf{w})^T\mathbf{v}_i$ form an orthonormal basis of the space of polynomials of order

d in $L^2_{\nu^d}(\mathbf{K})$. In the case where $M_d(\mathbf{z}^d)$ is singular, a regularization is introduced by considering the function

$$q_{\nu^d,\beta}(\mathbf{w}) = \mathbf{b}_d(\mathbf{w})^T(M_d(\mathbf{z}^d) + \beta\mathbf{I})^{-1}\mathbf{b}_d(\mathbf{w}) = \sum_{i=1}^{n_d}(\lambda_i + \beta)^{-1}(\mathbf{b}_d(\mathbf{w})^T\mathbf{p}_i)^2,$$

which turns out to be the inverse Christoffel-Darboux function of a measure $\nu^d + \beta\nu_0$, where ν_0 is the measure on $\mathbf{K}$ for which the monomials form an orthonormal family. Exploiting the fact that $q_{\nu^d+\beta\nu_0}$ tends to take low values on the graph of u, an approximation of u is defined by

$$f_{\beta,d}(t,\mathbf{x},\boldsymbol{\xi}) \in \arg\min_{y\in\mathbf{U}} q_{\nu^d+\beta\nu_0}(t,\mathbf{x},\boldsymbol{\xi},y).$$

Further information can be found in [28].

3.5.2 Statistical moments of variables of interest

Considering $\boldsymbol{\xi}$ as a random parameter, one may be interested in computing the expectation of some variable of interest $Q(\boldsymbol{\xi}) = F(u(\cdot,\cdot,\boldsymbol{\xi});\boldsymbol{\xi})$, where $F(\cdot,\boldsymbol{\xi})$ is a real-valued function taking as input time-space functions. In some particular situations, it is possible to directly obtain an estimation of this quantity from the moments $\mathbf{z}^d$. In particular, this is the case when

$$Q(\boldsymbol{\xi}) = \int_{\mathbf{T}\times\mathbf{X}} G(u(t,\mathbf{x},\boldsymbol{\xi}),t,\mathbf{x},\boldsymbol{\xi})dtd\mathbf{x},$$

with G is polynomial since then

$$\mathbb{E}(Q(\boldsymbol{\xi})) = \int_{\mathbf{T}\times\mathbf{X}\times\boldsymbol{\Xi}} G(\mathbf{w})d\nu(\mathbf{w}) \approx \ell_{\mathbf{z}^d}(G).$$

We may also be interested in obtaining statistical moments of the solution $u(t,\mathbf{x},\boldsymbol{\xi})$ at different points $(t,\mathbf{x})$, which is not a variable of interest in the above format. Of course, these quantities can be estimated from point-wise evaluations of u based on the technique presented in the previous section. However, an alternative approach is possible to estimate the statistical moments

$$\int_{\boldsymbol{\Xi}} u(t,\mathbf{x},\boldsymbol{\xi})^k d\rho(\boldsymbol{\xi}) := f_k(t,\mathbf{x}) \tag{34}$$

for all $(t,\mathbf{x}) \in \mathbf{T}\times\mathbf{X}$, from the the approximate moments $\mathbf{z}^d$ of the measure ν. We know that the measure ν can be disintegrated into its marginal $\lambda_{\mathbf{T}}\otimes\lambda_{\mathbf{X}}$ and its conditional measure $d\nu(\boldsymbol{\xi},y|t,\mathbf{x})$, such that $d\nu(t,\mathbf{x},\boldsymbol{\xi},y) = d\nu(\boldsymbol{\xi},y|t,\mathbf{x})dtd\mathbf{x}$. We assume that $f_k(t,\mathbf{x})$ takes values in a compact set $\mathbf{F} := [\underline{F},\overline{F}]$ which can be easily obtained in terms of $\mathbf{U}$ and k. We then let $\{\widetilde{g_j}\}_{j=1}^m$, $m\in\mathbb{N}$, be polynomials that describe the semi-algebraic compact set $\mathbf{T}\times\mathbf{X}\times\mathbf{F}$. Letting $\mathbf{z}$ be the sequence of moments of ν, we may notice that for all $\boldsymbol{\alpha} = (\alpha_1,\alpha_2) \in \mathbb{N}^{n+1}_{2d-k}$,

$$\mathbf{z}_{\alpha_1,\alpha_2,0,k} = \int_{\mathbf{T}}\int_{\mathbf{X}} t^{\alpha_1}\mathbf{x}^{\alpha_2} f_k(t,\mathbf{x})d\mathbf{x}dt = \int_{\mathbf{T}}\int_{\mathbf{X}}\int_{\mathbf{F}} t^{\alpha_1}\mathbf{x}^{\alpha_2} y\delta_{f_k(t,\mathbf{x})}(dy)d\mathbf{x}dt.$$

Our goal is then here to approximate the support of the measure $\delta_{f_k(t,\mathbf{x})}(dy)d\mathbf{x}dt$ from its moments $\boldsymbol{\omega}$ in order to recover the graph of $f^k(t,\mathbf{x})$. We are faced with

the issue that the information we have on the moments is incomplete, namely, we only have the moments $\omega_{\boldsymbol{\alpha},0}$ for $\boldsymbol{\alpha} \in \mathbb{N}^{n+1}_{2d}$ and $\omega_{\boldsymbol{\alpha},1}$ for $\boldsymbol{\alpha} \in \mathbb{N}^{n+1}_{2d-k}$. Following [18], we introduce the following finite-dimensional semi-definite programming (SDP) problems to recover the graph of $f_k(t, \mathbf{x})$ from incomplete moment information:

$$\begin{aligned} \inf_{\boldsymbol{\omega} \in \mathbb{R}^{(n+2)_d}} \quad & \mathrm{Tr}(M_d(\boldsymbol{\omega})) \\ \text{s.t.} \quad & \omega_{\boldsymbol{\alpha},0} = z^d_{\alpha_1,\alpha_2,0,0}, \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^{n+1}_{2d} \\ & \omega_{\boldsymbol{\alpha},1} = z^d_{\alpha_1,\alpha_2,0,k}, \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^{n+1}_{2d-k} \\ & M_{d-d_j}(\widetilde{g_j}\boldsymbol{\omega}) \succeq 0, \quad j = 0, ..., m, \end{aligned} \tag{35}$$

where $\mathrm{Tr}(M)$ denotes the trace of a matrix M. We recall that $M_d(\boldsymbol{\omega})$ denotes the moment matrix of $\boldsymbol{\omega}$. From this, we can compute the corresponding Christoffel-Darboux approximation of f_k, following the approach of the previous section, see [28, 18].

4 Numerical examples

For numerical illustration, we consider Burgers-type equations with parametrised initial condition or parametrised flux.

The choice of entropy pairs is important to ensure uniqueness of the solution. Implementing Kruzkhov's entropy pairs is possible (as seen in Section 3.1), but computationally heavy since it requires a reformulation with measures in higher dimension. It is known that the entropy $\eta(y) = y^2$ provides sufficient constraints to ensure uniqueness of the entropy solution for Burgers equation [9]. Then, instead of using Kruzkhov's pairs, we here rely on the following family of polynomial entropies:

$$\eta_l(y) = y^{2l}, \quad \forall l \in \mathbb{N}$$

and the corresponding polynomial functions $\mathbf{q}_l$. As an objective function, we choose the trace of the moment matrix (see discussion in section 3.4).

Numerical experiments are performed with the Matlab interface Gloptipoly3 [17].

In order to approximate the graph of solutions u, we use the method described in Section 3.5.1. Numerically, the optimization of the Christoffel function is achieved through a discretization of $\mathbf{T}$, $\mathbf{X}$, $\mathbf{\Xi}$ and $\mathbf{U}$ and the computation of the Christoffel function at each point of the grid.

We shall in the following denote by $\widetilde{u_d}$ the Christoffel-Darboux approximation of the solution using approximate moments from a degree d of the hierarchy, and by u the exact solution of our Riemann problem.

4.1 Riemann problem for the Burgers equation with parametrised initial condition

As a first example, we consider the classical one-dimensional Riemann problem (see e.g., [11]) for a Burgers equation, with a parameter-independent flux

$$\mathbf{f}(u) = \frac{1}{2}u^2.$$

and where we parametrise the initial position of the shock, taking

$$u_0(x,\xi) = \begin{cases} 1 \text{ if } x < \dfrac{1}{4}(\xi - 1), \\ 0 \text{ if } x \geq \dfrac{1}{4}(\xi - 1). \end{cases}$$

with a parameter ξ taking values in $\mathbf{\Xi} = [0,1]$. We know that the solution takes values in $\mathbf{U} = [0,1]$. The time-space window on which we consider the solution is $\mathbf{T} = [0, \frac{1}{2}]$ and $\mathbf{X} = [-\frac{1}{2}, \frac{1}{2}]$.

The unique solution is

$$u(t,x,\xi) = \begin{cases} 1 \text{ if } x < \dfrac{1}{4}(\xi - 1) + \dfrac{t}{2}, \\ 0 \text{ if } x \geq \dfrac{1}{4}(\xi - 1) + \dfrac{t}{2}, \end{cases} \tag{36}$$

Equipping $\mathbf{\Xi}$ with the Lebesgue measure on $[0,1]$, it yields the following statistical moments

$$\begin{aligned} f_k(t,x) &= \int_{\mathbf{\Xi}} u(t,x,\xi)^k d\rho(\xi) \\ &= 1 - \min(1, \max(0, 1 - 2t + 4x)) + 0^k \min(1, \max(0, 1 - 2t + 4x)), \end{aligned}$$

for all $k \in \mathbb{N}$, for all $(t,x) \in \mathbf{T} \times \mathbf{X}$. We may notice that in this simple case, f_k is independent on k for $k \geq 1$.

Retrieving the graph of the solution. Figure 1 shows the graphs of the approximate solution $\widetilde{u_d}(t,x,0)$ for $(t,x) \in \mathbf{T} \times \mathbf{X}$ (so that the shock is initially located at $x = -\frac{1}{4}$), with hierarchy's degree $d = 2, 5, 8$.

Figure 2 shows the graphs of the approximate solution $\widetilde{u_d}(t,x,1)$ for $(t,x) \in \mathbf{T} \times \mathbf{X}$ (so that the shock is initially located at $x = 0$), with hierarchy's degree $d = 2, 5, 8$.

We observe the same results as in [29], where discontinuities are very well resolved as early as $d = 5$.

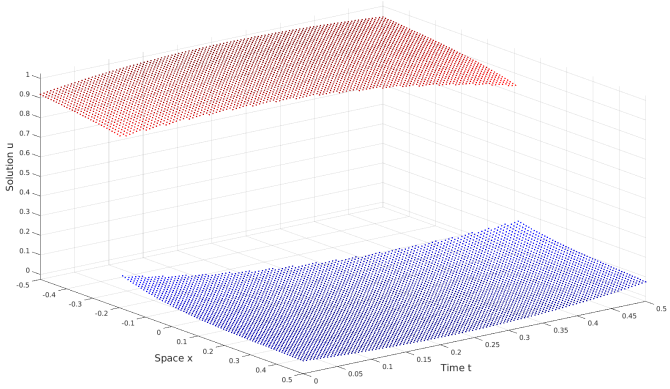
Error estimation. We choose to compute two different errors of our approximate solution.

$l^1(\mathbf{T} \times \mathbf{X} \times \mathbf{\Xi})$ error. We randomly pick 25 values in $\mathbf{\Xi}$, and consider 25 equidistant values in $\mathbf{T}$ and $\mathbf{X}$. We denote the test sets $\mathbf{\Xi}_e$, $\mathbf{T}_e$ and $\mathbf{X}_e$ respectively. We study the evolution of the relative l^1 error with respect to the degree d of the hierarchy. Namely, we are interested in

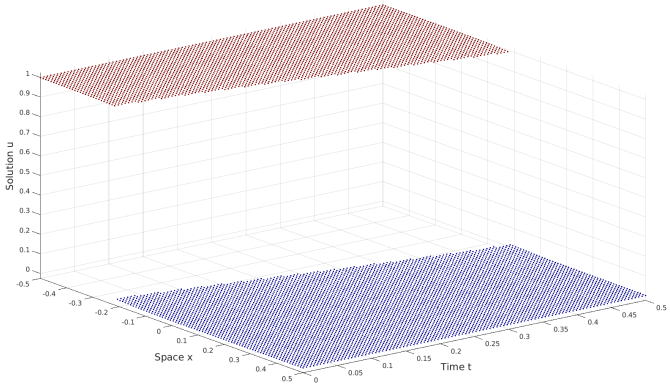
$$e_g(d) := \frac{\|u - \widetilde{u_d}\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e \times \mathbf{\Xi}_e)}}{\|u\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e \times \mathbf{\Xi}_e)}}.$$

The results are presented in Table 1.

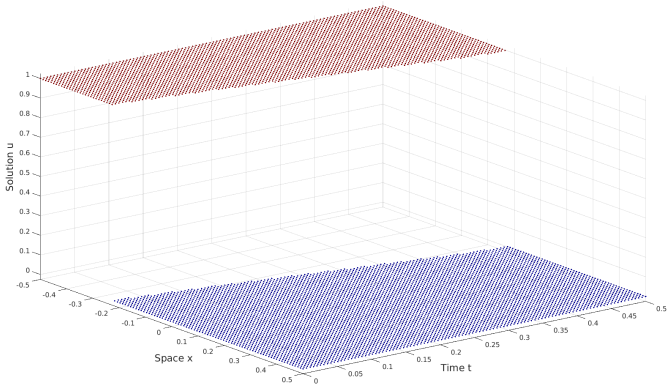
We observe a fast convergence of the error for small values of d. The convergence is not monotone and rather slow for high values of d.



(a) $d = 2$

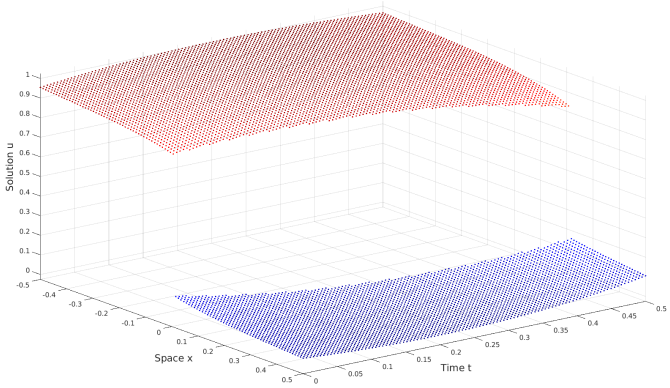


(b) $d = 5$

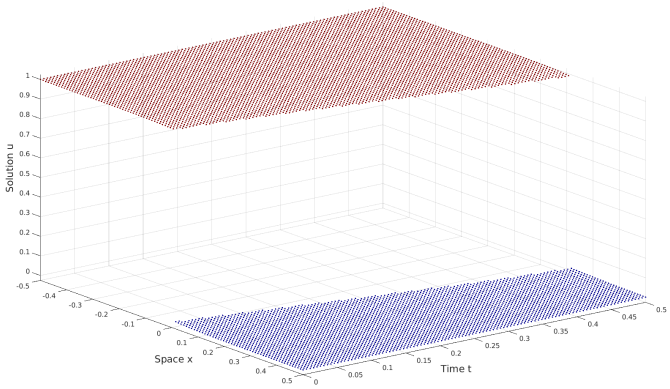


(c) $d = 8$

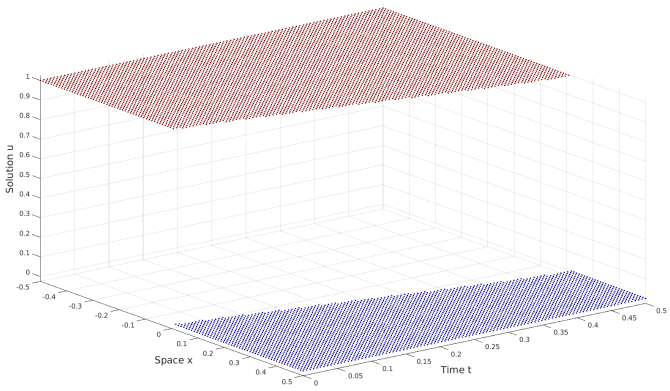
Figure 1: Graphs of the approximate solution $\widetilde{u_d}(t, x, 0)$ obtained with our approach for $d = 2, 5, 8$



(a) $d = 2$



(b) $d = 5$



(c) $d = 8$

Figure 2: Graphs of the approximate solution $\widetilde{u_d}(t, x, 0)$ obtained with our approach for $d = 2, 5, 8$

d	2	3	4	5	6	7	8
$e_g(d)$	0.0850	0.0267	0.0191	0.0168	0.0165	0.0167	0.0163

Table 1: $e_g(d)$ for $d = 2, \dots, 8$

$l^1(\mathbf{T} \times \mathbf{X})$ error for different parameter values. We consider four different values of the parameter $\xi \in \mathbf{\Xi}_e := (0, 0.2, 0.6, 1)$ (which correspond to a shock initially located at $x = -0.25$, $x = -0.2$, $x = -0.1$ and $x = 0$), and 100 equidistant points in $\mathbf{T}$ and $\mathbf{X}$, denoting the test sets $\mathbf{T}_e$ and $\mathbf{X}_e$ respectively. We then choose to study, for each $\xi_e \in \mathbf{\Xi}_e$, the evolution of the relative $l^1(\mathbf{T}_e \times \mathbf{X}_e)$ error with respect to the degree d of the hierarchy. We are thus interested in

$$e_{p_{\xi_e}}(d) := \frac{\|u(\cdot, \cdot, \xi_e) - \widetilde{u_d}(\cdot, \cdot, \xi_e)\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e)}}{\|u(\cdot, \cdot, \xi_e)\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e}},$$

for all $\xi_e \in \mathbf{\Xi}_e$. The results are presented in Table 2.

d	2	3	4	5	6	7	8
$e_{p0}(d)$	0.208	0.0616	0.0343	0.0314	0.0279	0.0276	0.0271
$e_{p0.2}(d)$	0.0971	0.0286	0.0218	0.0193	0.0176	0.0171	0.0182
$e_{p0.6}(d)$	0.0563	0.0207	0.0162	0.0162	0.0158	0.0161	0.0174
$e_{p1}(d)$	0.104	0.0407	0.0244	0.0229	0.0208	0.0194	0.0184

Table 2: $e_{p_{\xi_e}}(d)$ for $d = 2, \dots, 8$

We observe the same behaviour of the errors as in the previous paragraph.

Retrieving statistical moments of the solution. Denote $\mathbf{T}_e$ 100 equidistant points in $\mathbf{T}$ and $\mathbf{X}_e$ 100 equidistant points in $\mathbf{X}$. We want to approximate the expectation $f_1(x, t)$ of the solution for all $(t, x) \in \mathbf{T}_e \times \mathbf{X}_e$ following the method described in Section 3.5.2. Denoting $\widetilde{f_{1_d}}$ the approximated expected value of the solution for degree of relaxation d of the hierarchy, we want to compute the relative $l^1(\mathbf{T}_e \times \mathbf{X}_e)$ error of our approximation for $d = 2, \dots, 8$, namely, we are interested in

$$e_s(d) := \frac{\|f_1 - \widetilde{f_{1_d}}\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e)}}{\|f_1\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e)}},$$

for all $d = 2, \dots, 8$. The results are presented in Table 3.

d	2	3	4	5	6	7	8
$e_s(d)$	0.358	0.102	0.0557	0.0451	0.0484	0.0574	0.0637

Table 3: $e_s(d)$ for $d = 2, \dots, 8$

We note here the same phenomenon as for the errors presented above occurring, where the approximation rapidly improves as d rises until $d = 5$. The convergence is then rather slow and not monotone.

4.2 Riemann problem for the Burgers equation with parametrised flux

As a second illustration, we consider the classical one-dimensional Riemann problem (see e.g., [11]) for a Burgers equation, where we parametrise the flux of the equation. In particular, we choose the flux

$$f(u,\xi) = \frac{1}{4}(\xi+1)u^2,$$

with a parameter ξ taking values in $\mathbf{\Xi} = [0,1]$. The Riemann problem to this conservation law is a Cauchy problem with the following initial condition, piecewise constant with one point of discontinuity:

$$u_0(x) = \begin{cases} 1 \text{ if } x < 0, \\ 0 \text{ if } x \geq 0. \end{cases}$$

The solution is known to take values in $\mathbf{U} = [0,1]$. The time-space window on which we consider the solution is $\mathbf{T} = [0, \frac{1}{2}]$ and $\mathbf{X} = [-\frac{1}{2}, \frac{1}{2}]$.

The unique analytical solution corresponding to the initial condition is

$$u(t,x,\xi) = \begin{cases} 1 \text{ if } x < \dfrac{1}{4}(\xi+1)t, \\ 0 \text{ if } x \geq \dfrac{1}{4}(\xi+1)t, \end{cases} \tag{37}$$

We can note that the randomness in (36) was simply a translation of the solution, whereas, here, the phenomenon is non-linear, since the speed of the shock depends on ξ.

Providing $\mathbf{\Xi}$ with the Lebesgue measure on $[0,1]$, it comes that, for all $k \in \mathbb{N}$,

$$f_k(0,x) = u_0(x)^k,$$

for all $x \in \mathbf{X}$, and

$$f_k(t,x) = 1 - \min(1, \max(0, \frac{4x}{t} - 1)) + 0^k \min(1, \max(0, \frac{4x}{t} - 1)),$$

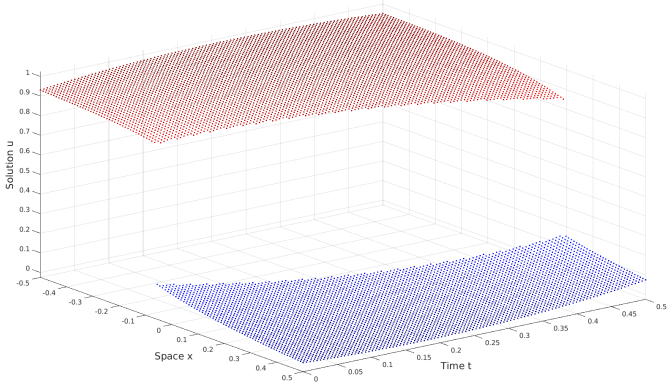
for all $(t,x) \in \mathbf{T} \times \mathbf{X}$. We may notice that in this simple case, for all $t > 0$, $f_k(t,\cdot)$ is independent on k for $k \geq 1$.

Retrieving the graph of the solution. Figure 3 shows the graphs of the approximate solution $\widetilde{u_d}(t,x,0)$ for $(t,x) \in \mathbf{T} \times \mathbf{X}$ (so that the speed of the shock is $\frac{1}{4}$), with hierarchy's degree $d = 2, 5, 8$.

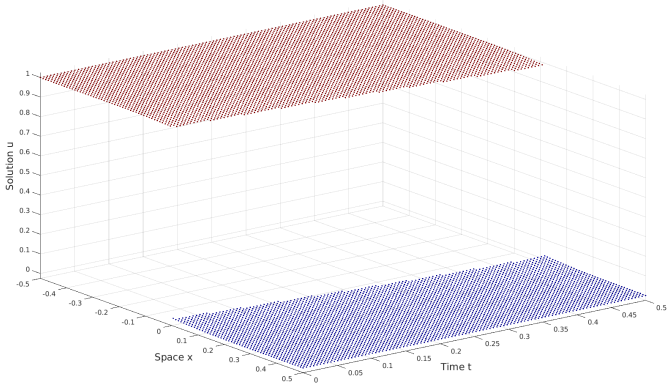
Figure 4 shows the graphs of the approximate solution $\widetilde{u_d}(t,x,1)$ for $(t,x) \in \mathbf{T} \times \mathbf{X}$ (so that the speed of the shock is $\frac{1}{2}$), with hierarchy's degree $d = 2, 5, 8$.

$l^1(\mathbf{T} \times \mathbf{X} \times \mathbf{\Xi})$ error of the method. We pick at random 25 values in $\mathbf{\Xi}$, and consider 25 equidistant values in $\mathbf{T}$ and $\mathbf{X}$. We denote the test sets $\mathbf{\Xi}_e$, $\mathbf{T}_e$ and $\mathbf{X}_e$ respectively. We study the evolution of the relative l^1 error with respect to the degree d of the hierarchy. Namely, we are interested in

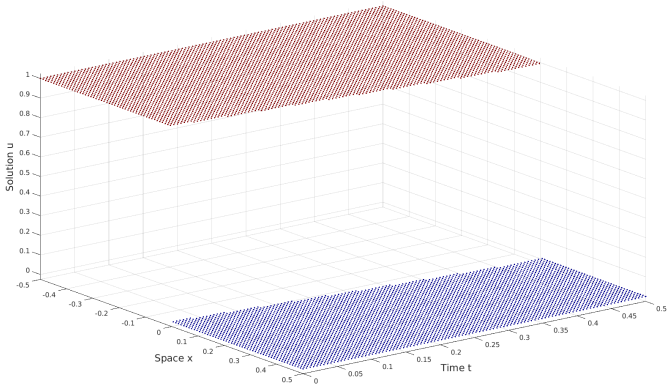
$$e_g(d) := \frac{\|u - \widetilde{u_d}\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e \times \mathbf{\Xi}_e)}}{\|u\|_{l^1(\mathbf{T}_e \times \mathbf{X}_e \times \mathbf{\Xi}_e)}}.$$



(a) $d = 2$

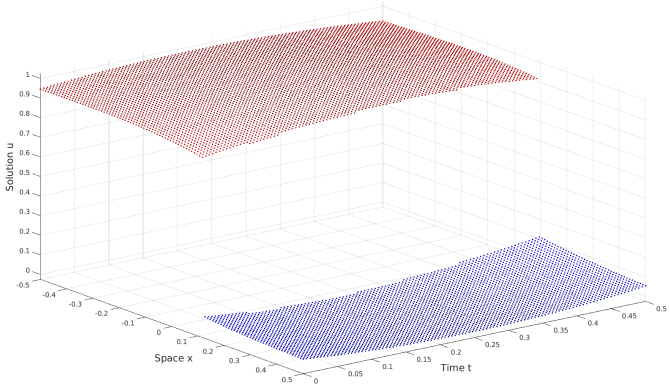


(b) $d = 5$

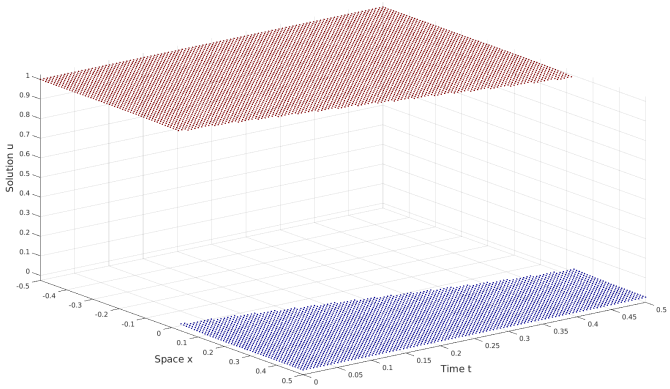


(c) $d = 8$

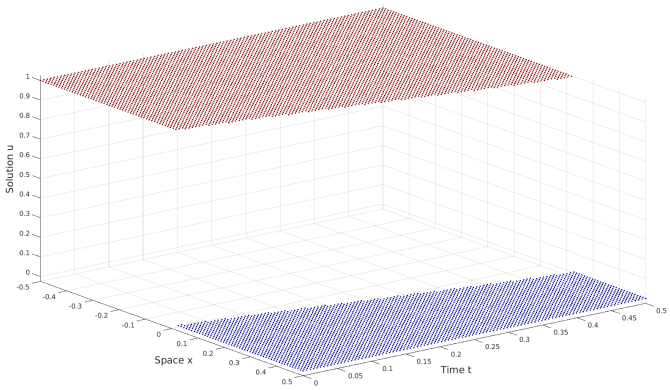
Figure 3: Graphs of the approximate solution $\widetilde{u_d}(t, x, 0)$ obtained with our approach for $d = 2, 5, 8$



(a) $d = 2$



(b) $d = 5$



(c) $d = 8$

Figure 4: Graphs of the approximate solution $\widetilde{u_d}(t, x, 1)$ obtained with our approach for $d = 2, 5, 8$

d	2	3	4	5	6	7	8
$e_g(d)$	0.0738	0.0285	0.0142	0.00772	0.00780	0.00818	0.00963

Table 4: $e_g(d)$ for $d = 2, \ldots, 8$

The results are shown in Table 4.

We note here the same phenomenon as for the errors presented above occurring, where the approximation improves as d rises until $d = 5$. Then the convergence is not monotone and rather slow.

Acknowledgments

The authors benefits from the support of the French National Research Agency within the project AIBY4 (Projet ANR-20-THIA-0011). The first and third authors thank the France 2030 framework programme Centre Henri Lebesgue ANR-11-LABX-0020-01 for its stimulating mathematical research programs.

Appendices

A Imposing marginal constraints of occupation measures

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

$$\int_{\mathbf{K}} t^{\alpha_1} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} d\nu(t, \mathbf{x}, \boldsymbol{\xi}, y) = \int_{\mathbf{T}\times\mathbf{X}\times\boldsymbol{\Xi}} t^{\alpha_1} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} dt d\mathbf{x} d\rho(\boldsymbol{\xi}), \tag{38}$$

for all $\boldsymbol{\alpha} \in \mathbb{N}^{n+p+1}$. In a similar manner, we impose the marginals of the time boundary measures to be products of a Dirac measure, a Lebesgue measure and ρ as follows: for all $\boldsymbol{\alpha} \in \mathbb{N}^{n+p+1}$,

$$\int_{\mathbf{K}} t^{\alpha_1} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} d\nu_0(t, \mathbf{x}, \boldsymbol{\xi}, y) = 0^{\alpha_1} \int_{\mathbf{X}\times\boldsymbol{\Xi}} \mathbf{x}^{\alpha_2} \boldsymbol{\xi}^{\alpha_3} d\mathbf{x} d\rho(\boldsymbol{\xi}), \tag{39}$$

$$\int_{\mathbf{K}} t^{\alpha_1} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} d\nu_T(t, \mathbf{x}, \boldsymbol{\xi}, y) = T^{\alpha_1} \int_{\mathbf{X}\times\boldsymbol{\Xi}} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} d\mathbf{x} d\rho(\boldsymbol{\xi}), \tag{40}$$

and

$$\int_{\mathbf{K}} t^{\alpha_1} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} d\nu_{B,i}(t, \mathbf{x}, \boldsymbol{\xi}, y) = \int_{\mathbf{T}\times\Gamma_{B,i}\times\boldsymbol{\Xi}} t^{\alpha_1} \mathbf{x}^{\boldsymbol{\alpha}_2} \boldsymbol{\xi}^{\boldsymbol{\alpha}_3} dt d\mathbf{x} d\rho(\boldsymbol{\xi}), \tag{41}$$

for all $1 \leq i \leq n$ and $B \in \{L, R\}$.

B Split measures and corresponding moments constraints

In addition to split measures ϑ_T^+ and ϑ_T^- associated with ν, we introduce the time boundary measures ϑ_0^+, ϑ_0^-, ϑ_T^+ and ϑ_T^- , which are defined as

$$d\vartheta_0^{\pm}(t,\mathbf{x},\boldsymbol{\xi},y,v) := \mathbb{1}_{\{u\in\mathbf{U}:\pm(y-u)\geq 0\}}(v)d\nu_0(t,\mathbf{x},\boldsymbol{\xi},y)dv, \tag{42a}$$

$$d\vartheta_T^{\pm}(t,\mathbf{x},\boldsymbol{\xi},y,v) := \mathbb{1}_{\{u\in\mathbf{U}:\pm(y-u)\geq 0\}}(v)d\nu_T(t,\mathbf{x},\boldsymbol{\xi},y)dv, \tag{42b}$$

with supports

$$\mathbf{K}_0^{\pm} := \mathrm{supp}(\vartheta_0^+) = \{(t,\mathbf{x},\boldsymbol{\xi},y,v)\in\mathbf{K}_0\times\mathbf{U} : \pm(y-v)\geq 0\}, \tag{43}$$

$$\mathbf{K}_T^{\pm} := \mathrm{supp}(\vartheta_T^+) = \{(t,\mathbf{x},\boldsymbol{\xi},y,v)\in\mathbf{K}_T\times\mathbf{U} : \pm(y-v)\geq 0\}, \tag{44}$$

respectively. We only introduce the space boundary measures $(\vartheta_{L_i}^+)_{i=1}^n$, $(\vartheta_{L_i}^-)_{i=1}^n$, $(\vartheta_{R_i}^+)_{i=1}^n$ and $(\vartheta_{R_i}^-)_{i=1}^n$, defined as

$$d\vartheta_{B,i}^{\pm}(t,\mathbf{x},\boldsymbol{\xi},y,v) := \mathbf{1}_{\{u\in\mathbf{U}:\pm(y-u)\geq 0\}}(v)d\nu_{L_i}(t,\mathbf{x},\boldsymbol{\xi},y)dv, \tag{45}$$

for $B\in\{L,R\}$ and $1\leq i\leq n$, with supports

$$\mathbf{K}_{B,i}^{\pm} := \mathrm{supp}(\vartheta_{B,i}^{\pm}) = \{(t,\mathbf{x},\boldsymbol{\xi},y,v)\in\mathbf{K}_{B,i}\times\mathbf{U} : \pm(y-v)\geq 0\}. \tag{46}$$

The relation between ν and split measures ϑ^+ and ϑ^- is imposed through moment constraints

$$\int_{\mathbf{K}\times\mathbf{U}} \mathbf{w}^{\boldsymbol{\alpha}} v^{\beta}(d\vartheta^+ + \vartheta^-)(\mathbf{w},v) = \int_{\mathbf{K}\times\mathbf{U}} \mathbf{w}^{\boldsymbol{\alpha}} v^{\beta} d\nu(\mathbf{w})dv, \tag{47}$$

for all $\boldsymbol{\alpha}\in\mathbb{N}^{n+p+2}$ and for all $\beta\in\mathbb{N}$. Similar conditions are imposed between time and boundary measures and their corresponding split measures.

References

[1] R. Abgrall and S. Mishra. Chapter 19 - uncertainty quantification for hyperbolic systems of conservation laws. In Rémi Abgrall and Chi-Wang Shu, editors, Handbook of Numerical Methods for Hyperbolic Problems, volume 18 of Handbook of Numerical Analysis, pages 507–544. Elsevier, 2017.

[2] C. D. Aliprantis and K. C. Border. Infinite Dimensional Analysis. Springer Berlin, Heidelberg, 2006.

[3] Ja. Badwaik, C. Klingenberg, N. H. Risebro, and A. M. Ruf. Multilevel monte carlo finite volume methods for random conservation laws with discontinuous flux. ESAIM: M2AN, 55(3):1039–1065, 2021.

[4] C. Bardos, A.-Y. Le Roux, and J. C. Nedelec. First order quasilinear equations with boundary conditions. Communications in Partial Differential Equations, 4:1017–1034, 1979.

[5] H. Bijl, D. Lucor, S. Mishra, and C. Schwab. Uncertainty Quantification in Computational Fluid Dynamics. Lecture Notes in Computational Science and Engineering. Springer International Publishing, 2013.

[6] A.-C. Boulanger, P. Moireau, B. Perthame, and J. Sainte-Marie. Data Assimilation for hyperbolic conservation laws. A Luenberger observer approach based on a kinetic description. Communications in Mathematical Sciences, 13(3):587 – 622, March 2015.

[7] C. Chalons, R. Duvigneau, and C. Fiorini. Sensitivity analysis and numerical diffusion effects for hyperbolic PDE systems with discontinuous solutions. the case of barotropic euler equations in lagrangian coordinates. SIAM Journal on Scientific Computing, 40(6):A3955–A3981, 2018.

[8] C. M. Dafermos. Hyperbolic Conservation Laws in Continuum Physics; 3rd ed. Grundlehren der mathematischen Wissenschaften: a series of comprehensive studies in mathematics. Springer, Dordrecht, 2010.

[9] C. De Lellis, F. Otto, and M. Westdickenberg. Minimal entropy conditions for burgers equation. Quarterly of Applied Mathematics, 62(4):687–700, 2004.

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

[11] L. C. Evans. Partial differential equations and Monge-Kantorovich mass transfer. Current developments in mathematics, 1997(1):65–126, 1997. Publisher: International Press of Boston.

[12] R. Eymard, T. Gallouët, and R. Herbin. Finite Volume Methods. In J. L. Lions and Philippe Ciarlet, editors, Solution of Equation in $\mathbb R^n$ (Part 3), Techniques of Scientific Computing (Part 3), volume 7 of Handbook of Numerical Analysis, pages 713–1020. Elsevier, 2000.

[13] J. Giesselmann, F. Meyer, and C. Rohde. A posteriori error analysis and adaptive non-intrusive numerical schemes for systems of random conservation laws. BIT Numerical Mathematics, 60, 03 2020.

[14] E. Godlewski and P-A Raviart. Hyperbolic Systems Of Conservation Laws. Ellipses, 1991.

[15] S. Grundel and M. Herty. Model-order reduction for hyperbolic relaxation systems. International Journal of Nonlinear Sciences and Numerical Simulation, 2022.

[16] D. Henrion, M. Infusino, S. Kuhlmann, and V. Vinnikov. Infinite-dimensional moment-sos hierarchy for nonlinear partial differential equations. arXiv preprint arXiv:2305.18768, 2023.

[17] D. Henrion, J. B. Lasserre, and J. Lofberg. Gloptipoly 3: moments, optimization and semidefinite programming, 2007.

[18] D. Henrion and J.B. Lasserre. Graph Recovery from Incomplete Moment Information. Constructive Approximation, 56:165–187, 2022.

[19] H. Holden and N.H. Risebro. Front Tracking for Hyperbolic Conservation Laws. Applied Mathematical Sciences. Springer Berlin Heidelberg, 2015.

[20] M. Korda, D. Henrion, and J. B. Lasserre. Moments and convex optimization for analysis and control of nonlinear partial differential equations. In Elsevier, editor, Handbook of Numerical Analysis, volume 23, pages 339–366. Elsevier, 2022.

[21] S. G. Krupa and A. F. Vasseur. On uniqueness of solutions to conservation laws verifying a single entropy condition. Journal of Hyperbolic Differential Equations, 16(01):157–191, mar 2019.

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

[23] F. Laakmann and P. Petersen. Efficient approximation of solutions of parametric linear transport equations by relu dnns. Advances in Computational Mathematics, 47, 02 2021.

[24] J. B. Lasserre. Moments, Positive Polynomials and Their Applications. Imperial College Press, 2009.

[25] J. B. Lasserre. The Christoffel-Darboux Kernel for Data Analysis. In 23ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, Villeurbanne - Lyon, France, February 2022. INSA Lyon.

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

[27] R. J. LeVeque. Numerical methods for conservation laws (2. ed.). Lectures in mathematics. Birkhäuser, 1992.

[28] S. Marx, E. Pauwels, T. Weisser, D. Henrion, and J. B. Lasserre. Semi-algebraic approximation using Christoffel-Darboux kernel. Constructive Approximation, April 2021. Publisher: Springer Verlag.

[29] S. Marx, T. Weisser, D. Henrion, and J. B. Lasserre. A moment approach for entropy solutions to nonlinear hyperbolic PDEs. Mathematical Control and Related Fields, 10(1):113–140, January 2020. Publisher: AIMS.

[30] S. Mishra, N. H. Risebro, C. Schwab, and S. Tokareva. Numerical Solution of Scalar Conservation Laws with Random Flux Functions. SIAM/ASA Journal on Uncertainty Quantification, 4(1):552–591, 2016. _eprint: https://doi.org/10.1137/120896967.

[31] S. Mishra and C. Schwab. Sparse tensor multi-level monte carlo finite volume methods for hyperbolic conservation laws with random initial data. Mathematics of Computation, 81(280):1979–2018, 2012. Publisher: American Mathematical Society.

[32] J. Nečas, J. Málek, M. Rokyta, and M. Ružička. Weak and measure-valued solutions to evolutionary PDEs, volume 13 of Appl. Math. Math. Comput. London: Chapman & Hall, 1996.

[33] F. Otto. Initial-boundary value problem for a scalar conservation law. Comptes Rendus de l'Académie des Sciences. Série I, 322(8):729–734, 1996.

[34] G. Poëtte, B. Després, and D. Lucor. Uncertainty quantification for systems of conservation laws. Journal of Computational Physics, 228(7):2443–2467, 2009.

[35] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. SIAM Review, 52(3):471–501, 2010. _eprint: https://doi.org/10.1137/070697835.

[36] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann. The shifted proper orthogonal decomposition: A mode decomposition for multiple transport phenomena. SIAM Journal on Scientific Computing, 40(3):A1322–A1344, 2018.

[37] M. Tacchi. Convergence of Lasserre's hierarchy: the general case. Optimization Letters, 16:1–19, 06 2021.

[38] G. B. Whitham. Linear and nonlinear waves, volume 42. John Wiley & Sons, 2011.

[39] X. Zhong and C.-W. Shu. Entropy stable galerkin methods with suitable quadrature rules for hyperbolic systems with random inputs. Journal of Scientific Computing, 92, 06 2022.