

Monte-Carlo Finite-Volume Methods in Uncertainty Quantification for Hyperbolic Conservation Laws

S. Mishra and Ch. Schwab

Research Report No. 2017-38

August 2017

Latest revision: May 2018

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Monte-Carlo Finite-Volume Methods in Uncertainty Quantification for Hyperbolic Conservation Laws

Siddhartha Mishra and Christoph Schwab

Abstract

We consider hyperbolic systems of conservation laws and review developments in the general area of computational uncertainty quantification (UQ) for these equations. We focus on non-intrusive sampling methods of the Monte-Carlo (MC) and Multi-level Monte-Carlo (MLMC) type. The modeling of uncertainty, within the framework of random fields and random entropy solutions, is discussed. We also describe (ML)MC finite volume methods and present the underlying error bounds and complexity estimates. Based on these bounds, and numerical experiments, we illustrate the gain in efficiency resulting from the use of MLMC methods in this context. Recent progress in the mathematical UQ frameworks of measure-valued and statistical solutions is briefly presented, with comprehensive literature survey.

1 Introduction

Systems of conservation laws are nonlinear partial differential equations of the form

$$\partial_t \mathbf{U} + \nabla_x \cdot \mathbf{F}(\mathbf{c}(x, t), \mathbf{U}) = 0, \quad (1a)$$

$$\mathbf{U}(x, 0) = \mathbf{U}_0(x). \quad (1b)$$

Here, the unknown $\mathbf{U} = \mathbf{U}(x, t) : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^N$ is the vector of *conserved variables*, $\mathbf{F} = (\mathbf{F}^1, \dots, \mathbf{F}^d) : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^{N \times d}$ is the *flux function* and $\mathbf{c} = (\mathbf{c}^1, \dots, \mathbf{c}^d) : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^{N \times d}$ is a spatio-temporal coefficient. We denote $\mathbb{R}_+ := [0, \infty)$. Here, \mathbf{U}_0 denotes the prescribed initial data.

Siddhartha Mishra · Christoph Schwab

Seminar for Applied Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland e-mail: smishra@sam.math.ethz.ch, e-mail: christoph.schwab@sam.math.ethz.ch

In bounded domains, the system (1) needs to be supplemented with suitable boundary conditions.

The system (1) is termed *hyperbolic* if the flux Jacobian matrix has real eigenvalues [11]. Hyperbolic systems of conservation laws arise in a wide variety of models in physics and engineering and we refer to [11] for a wide range of examples. Solutions of (1) can develop discontinuities in finite time, even for smooth initial data (see again [11] and the references there). Therefore, solutions of (1) are weak solutions in that $\mathbf{U} \in (L^1_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}_+))^N$ is required to satisfy the integral identity

$$\int_{R_+} \int_{R^d} \left(\mathbf{U} \varphi_t + \sum_{j=1}^d \mathbf{F}^j(\mathbf{c}^j, \mathbf{U}) \varphi_{x_j} \right) dx dt + \int_{R^d} \mathbf{U}_0(x) \varphi(x, 0) dx = 0, \quad (2)$$

for all test functions $\varphi \in C_0^1(\mathbb{R}_+ \times \mathbb{R}^d)$. It is well known that weak solutions are not necessarily unique [11]. Additional admissibility criteria or *entropy conditions* are necessary to obtain uniqueness. In space dimension $d > 1$, rigorous existence and uniqueness results for conservation (balance) laws and for generic initial data are available only for the scalar case, i.e. in the case $N = 1$.

1.1 Numerical methods

Numerical methods for the solution of (2) comprise Finite Difference (FD), Finite Volume (FV) and Discontinuous Galerkin (DG) methods. We refer to the textbooks [42, 28] and the references there.

Within the popular FV framework [28], the cell averages of the unknown are updated in time in terms of numerical fluxes across cell interfaces. These numerical fluxes are often obtained by the (approximate) solutions of Riemann problems in the direction normal to the cell interface. Higher order spatial accuracy is obtained by reconstructing cell averages in terms of non-oscillatory piecewise polynomial functions, within the TVD [42], ENO [33] and WENO [6] procedures or using Discontinuous Galerkin methods (see, e.g. [8]). Higher order temporal accuracy is achieved by employing strong stability preserving Runge-Kutta methods [30]. Space-time DG-discretizations can also be employed for High-order spatio-temporal accuracy [34].

1.2 Aims and scope

Any numerical scheme approximating (1) requires the initial data \mathbf{U}_0 , the coefficients \mathbf{c} and the flux function \mathbf{F} , as well as suitable boundary conditions, as inputs. However, in practice, these inputs are obtained by measurements (observations). Moreover, measurements cannot be precise and always involve some degree of uncertainty. Input

uncertainty for (1) implies, upon uncertainty propagation, corresponding uncertainty in the solution.

The modeling, mathematical analysis, numerical approximation and numerical quantification of solution uncertainty, given experimental data, comprise the discipline of uncertainty quantification (UQ).

One aim in this article is to survey computational methods for the efficient, computational UQ for nonlinear, hyperbolic conservation laws with random inputs, and to provide some indications on the numerical analysis of UQ methods for these equations. Our focus in this article is on *non-intrusive computational methods* and their implementation, and on the mathematical analysis of their computational complexity. In our presentation, we emphasize broad applicability for a large class of conservation laws, rather than problem-specific, optimal results.

Our motivation for this focus on non-intrusive methods is as follows: first, non-intrusive methods afford trivial integration of existing, deterministic numerical solvers of instances of forward problems, and are, therefore, popular in computational UQ in science and engineering. Second and as mentioned earlier, nonlinear hyperbolic conservation laws are well known to exhibit *solutions of very low regularity in physical space*, due to *shock formation* even for smooth input data (initial and boundary data, as well as flux functions). Third, *hyperbolicity implies finite speed of propagation* which, in the context of UQ for conservation laws with parametric input uncertainty, implies propagation of singular supports into the domain of parameters that describe the uncertain inputs of the system. The presence of, in general, moving singular supports propagating along characteristics in parametric families of weak solutions precludes high convergence rates of “smooth” computational methods, such as generalized polynomial chaos, PCA etc. for this class of computational UQ problems (we mention, however, that even in the absence of viscosity, there are regularizing effects due to averaging; cases in point are the so-called “transport collapse” regularizations in averaging lemmas (see, e.g. [43, 44] and the references there) or due to statistical ensemble averaging of random entropy solutions (see, e.g. [57])).

We therefore focus in the present survey on *sampling methods* of Monte-Carlo (MC for short) and of Multi-Level MC (MLMC for short) type, as well as on stochastic collocation methods. These methods have in common that their computational realization is based on existing numerical conservation law solvers, for example the finite volume (FV) or discontinuous Galerkin (DG) type, without any modification; this implies, in particular, that existing discretization error bounds for these methods, e.g. from [7, 15, 39, 42] and the references there, can be used for an error analysis of non-intrusive computational UQ for hyperbolic conservation laws.

In contrast, *intrusive computational methods* will require, as a rule, some form of reformulation of the conservation law prior to discretization and entail, usually, significant refactoring resp. redesign of numerical solvers. We refer to, e.g., [61, 24] and references therein for examples of this type, where the so-called *stochastic Galerkin* methodology has been employed and was shown to require significant modifications of numerical schemes as well as of actual, numerical solver.

Given our focus on non-intrusive UQ methods of the MC and Multi-level MC type, we structure this survey as follows: in the first part, we will focus on the very specific problem of UQ for *scalar* conservation laws with random initial data. Here, we describe data and solution uncertainty in terms of random fields and within the framework of random entropy solutions [47]. This is feasible as the underlying deterministic solution operator is well-defined and forms a non-expansive (in time) semi-group on $L^1(\mathbb{R}^d)$. We will formulate both the MC and MLMC methods and combine them with a FV space-time discretization to obtain rigorous convergence rates for the (ML)MC-FV scheme and demonstrate that the MLMC-FV method is significantly more efficient (computationally) than the MC-FV method.

Next, we extend the (ML)MC-FV schemes for UQ of systems of conservation laws with random inputs. Here, the underlying deterministic problem may be ill-posed within the class of entropy solutions [14]. Consequently, the notion of random entropy solutions may not be well-defined. Moreover, there is no rigorous convergence result for the underlying deterministic FV (or any other) discretization frameworks. Hence, we postulate convergence and obtain the corresponding error (and complexity) estimates for the (ML)MC-FV methods. Although this combination is seen to work well in practice, recent results [16, 45, 19] have demonstrated the limitations of this framework. Instead, novel solution concepts such as those of *entropy measure valued solutions* [16, 19] and *statistical solutions* [17] and have been proposed and analyzed. We will conclude with a brief review of these concepts.

2 Preliminaries

2.1 Random variables in Banach spaces

Our mathematical formulation of scalar conservation laws with random inputs will use the concept of random fields i.e. random variables taking values in function spaces. We recapitulate basic concepts as presented, for example, in [10, Chap1].

Let E be a Banach space, and let (Ω, \mathcal{F}) be a measurable space, with the set Ω of elementary events, and with \mathcal{F} a corresponding σ -algebra. An E -valued random variable (or random variable taking values in E) is any mapping $X : \Omega \rightarrow E$ such that the set $\{\omega \in \Omega : X(\omega) \in A\} = \{X \in A\} \in \mathcal{F}$ for any $A \in \mathcal{G}$, i.e. such that X is a \mathcal{G} -measurable mapping from Ω into E . Here, (E, \mathcal{G}) denotes a measurable space on the Banach space E .

For a Banach space E , we denote the Borel σ -field $\mathcal{B}(E)$. Then, $(E, \mathcal{B}(E))$ is a measurable space and random variables taking values in E i.e. maps $X : \Omega \rightarrow E$ are $(\mathcal{F}, \mathcal{B}(E))$ measurable. For a separable Banach space E with norm $\|\cdot\|_E$ and (topological) dual E^* , $\mathcal{B}(E)$ is the smallest σ -field of subsets of E containing all sets

$$\{x \in E : \varphi(x) \leq \alpha\}, \quad \varphi \in E^*, \quad \alpha \in \mathbb{R}.$$

For a separable Banach space, $X : \Omega \rightarrow E$ is an E -valued random variable iff for every $\varphi \in E^*$, $\omega \mapsto \varphi(X(\omega)) \in \mathbb{R}^1$ is an \mathbb{R}^1 -valued random variable: for any RV $X : \Omega \rightarrow E$ on (Ω, \mathcal{F}) which takes values in E , the mapping $\Omega \ni \omega \mapsto \|X(\omega)\|_E \in \mathbb{R}^1$ is (strongly) measurable. For more details and proofs, we refer to [10] or to [51].

The strongly measurable mapping $X : \Omega \rightarrow E$ is *Bochner integrable* if, for any probability measure \mathbb{P} on the measurable space (Ω, \mathcal{F}) ,

$$\int_{\Omega} \|X(\omega)\|_E \mathbb{P}(d\omega) < \infty. \quad (3)$$

A probability measure \mathbb{P} on (Ω, \mathcal{F}) is a σ -additive set function from Ω into $[0, 1]$ such that $\mathbb{P}(\Omega) = 1$; the triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called probability space. We shall always assume, unless explicitly stated, that $(\Omega, \mathcal{F}, \mathbb{P})$ is complete.

An E -valued RV is called *simple* if it can assume only finitely many values. A simple RV X , taking values in E , has the explicit form (with χ_A denoting the indicator function of $A \in \mathcal{F}$)

$$X = \sum_{i=1}^N x_i \chi_{A_i}, \quad A_i \in \mathcal{F}, \quad x_i \in E, \quad N < \infty. \quad (4)$$

For simple RVs X taking values in E and for any $B \in \mathcal{F}$,

$$\int_B X(\omega) \mathbb{P}(d\omega) = \int_B X d\mathbb{P} := \sum_{i=1}^N x_i \mathbb{P}(A_i \cap B). \quad (5)$$

For such $X(\cdot)$ and for all $B \in \mathcal{F}$,

$$\left\| \int_B X(\omega) \mathbb{P}(d\omega) \right\|_E \leq \int_B \|X(\omega)\|_E \mathbb{P}(d\omega). \quad (6)$$

For any random variable $X : \Omega \rightarrow E$ which is Bochner integrable, there exists a sequence $\{X_m\}_{m \in \mathbb{N}}$ of simple random variables such that, for all $\omega \in \Omega$, $\|X(\omega) - X_m(\omega)\|_E \rightarrow 0$ as $m \rightarrow \infty$. Therefore, (5) and (6) extend in the usual fashion by continuity to any E -valued random variable. We denote the Bochner-integral

$$\int_{\Omega} X(\omega) \mathbb{P}(d\omega) = \lim_{m \rightarrow \infty} \int_{\Omega} X_m(\omega) \mathbb{P}(d\omega) \in E \quad (7)$$

by $\mathbb{E}[X]$ (“expectation” of X).

We introduce for $1 \leq p \leq \infty$ Bochner spaces of p -summable random variables X taking values in the Banach-space E . The set $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ comprises all (equivalence classes of) integrable, E -valued random variables X . It is a Banach space if equipped with the norm

$$\|X\|_{L^1(\Omega; E)} := \mathbb{E}(\|X\|_E) = \int_{\Omega} \|X(\omega)\|_E \mathbb{P}(d\omega). \quad (8)$$

Define $L^p(\Omega, \mathcal{F}, \mathbb{P}; E)$ for $1 \leq p < \infty$ as the set of p -summable random variables taking values E . With the norm

$$\|X\|_{L^p(\Omega; E)} := (\mathbb{E}(\|X\|_E^p))^{1/p}, \quad 1 \leq p < \infty \quad (9)$$

$L^p(\Omega, \mathcal{F}, \mathbb{P}; E)$ becomes a Banach space. For $p = \infty$, we denote by $L^\infty(\Omega, \mathcal{F}, \mathbb{P}; E)$ the set of all E -valued random variables which are essentially bounded. This set is a Banach space equipped with the norm

$$\|X\|_{L^\infty(\Omega; E)} := \operatorname{ess\,sup}_{\omega \in \Omega} \|X(\omega)\|_E. \quad (10)$$

If $T < \infty$ and $\Omega = [0, T]$, $\mathcal{F} = \mathcal{B}([0, T])$, we write $L^p([0, T]; E)$. Note that for any separable Banach-space E , and for any $r \geq p \geq 1$,

$$L^r(0, T; E), C^0([0, T]; E) \in \mathcal{B}(L^p(0, T; E)). \quad (11)$$

We conclude the section of preliminaries with a criterion for strong measurability.

Lemma 1. [51, Corollary 1.13] *Let E_1 and E_2 be Banach spaces, and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. If $f : \Omega \rightarrow E_1$ is strongly Bochner measurable, and if $\phi : E_1 \rightarrow E_2$ is continuous, then the composition $\phi \circ f : \Omega \rightarrow E_2$ is strongly Bochner measurable.*

2.2 Monte-Carlo (MC) Sampling in Banach Spaces

Let $\widehat{Y}_i : \Omega \rightarrow F$, $i = 1, \dots, M$, be independent identically distributed (“iid” for short) random variables taking values in the Banach space E . We let

$$E_M[Y^{(k)}] := \frac{1}{M} \sum_{i=1}^M \widehat{Y}_i^{(k)},$$

be the Monte Carlo estimator for $\mathbb{E}[Y^{(k)}]$. A computable estimate $E_M[Y^{(k)}]$ for the k -th moment $\mathcal{M}^k(Y)$ of Y will converge in E as $M \rightarrow \infty$ at a rate which depends on the integrability of Y . Specifically, at the rate of convergence (in terms of M) of

$$\mathbb{E} \left[\left\| \mathbb{E}[Y^{(k)}] - E_M[Y^{(k)}] \right\|_E^p \right]^{1/p}$$

to zero as $M \rightarrow \infty$ for some $1 \leq p < \infty$. If E is a finite dimensional space or a Hilbert space, and if $X \in L^{2k}(\Omega; X)$, the so-called *mean square error* (MSE) is bounded as

$$\mathbb{E} \left[\left\| \mathbb{E}[Y^{(k)}] - E_M[Y^{(k)}] \right\|_E^2 \right] \leq \frac{C}{M} \operatorname{Var}[Y^{(k)}] \quad (12)$$

using the independence of the samples \widehat{Y}_i .

For general Banach spaces E , the convergence rate depends on the *type* of the Banach space, which is defined as follows, [38, Page 246].

Definition 1. Let $1 \leq p \leq \infty$ and $Z_j, j \in \mathbb{N}$ a sequence of Bernoulli-Rademacher random variables. A Banach space E is said to be of *type* p if there is a *type constant* $C > 0$ such that for all finite sequences $(x_j)_{j=1}^N \subset E, N \in \mathbb{N}$,

$$\left\| \sum_{j=1}^N Z_j x_j \right\|_E \leq C \left(\sum_{j=1}^N \|x_j\|_E^p \right)^{1/p}.$$

By the triangle inequality, every Banach space has type 1. Hilbert spaces have type 2. One can show that the L^p -spaces have type $\min\{p, 2\}$ for $1 \leq p < \infty$, [38].

One has the following result from [36], [38, Proposition 9.11] and [9, Section 4] for Banach spaces of type p .

Proposition 1. [36] *Let E be a Banach space of type p with a type constant C_t . Then, for every finite sequence $(Y_j)_{j=1}^N$ of independent mean zero random variables in $L^p(\Omega)$, one has the bound*

$$\mathbb{E} \left[\left\| \sum_{j=1}^N Y_j \right\|_E^p \right] \leq (2C_t)^p \sum_{j=1}^N \mathbb{E} [\|Y_j\|_E^p].$$

Corollary 1. *Let E be a Banach space of type $p \in [1, 2]$ with type constant C_t . Then for every finite sequence $(Y_j)_{j=1}^N$ of iid mean zero random variables with $Y_j(\omega) \sim Y(\omega)$ in $L^p(\Omega)$,*

$$\mathbb{E} \left[\left\| E_M[Y^{(k)}] \right\|_E^p \right] = \mathbb{E} \left[\left\| \frac{1}{N} \sum_{j=1}^N Y_j^{(k)} \right\|_E^p \right] \leq (2C_t)^p N^{1-p} \mathbb{E} [\|Y^{(k)}\|_E^p].$$

Remark 1. The complexity of MC methods with respect to the type parameter p of the function space E has been investigated in [12, 13] and the references there. The relevance of Proposition 1 for MC methods applied to scalar conservation laws is due to $L^1(\mathbb{R}^d)$ being crucial for the error- and well-posedness analysis of the underlying FV schemes for these problems. The space $L^1(\mathbb{R}^d)$ being a Banach space of type 1, will a priori not allow for convergence rate bounds in MLMC-FV discretizations, as was incorrectly stipulated in [47], and also in related work [54] on combining MLMC discretization with the front-tracking algorithm.

Instead, and as pointed out in [46], lower rates of convergence of FV discretizations in the stronger norms $L^p(\mathbb{R}^d)$ with $p > 1$, with the space $L^p(\mathbb{R}^d)$ being of type $\min\{p, 2\}$, lead to convergence and error vs. work analysis of (ML)MC-FV methods for scalar conservation laws, which will be described in detail subsequently.

3 Scalar conservation laws with random initial data

We begin with a review of classical results on deterministic scalar conservation laws (SCLs). We also review random entropy solutions for SCLs with random initial data from [47], and in particular existence and uniqueness of a random entropy solution with finite second moments. Let us mention that SCLs with random input data have received considerable attention in the context of numerical methods for uncertainty quantification; we only mention [22, ?, 63].

We also mention considerable activity on enlarging the class of admissible random flux functions, in particular by so-called “rough path” calculus and the kinetic (re)formulation of the SCL (14) - (15) rather than the Kruřkov theory; we refer to [22, 44, 43, 26] and the references there.

3.1 Deterministic scalar hyperbolic conservation laws

In-order to present the basic ideas in a simple setting, we consider the Cauchy problem for scalar conservation laws (SCL) i.e, (1) with $N = 1$ and with a spatially homogeneous deterministic flux function $f(u)$. Then, (1) can be written as

$$\frac{\partial u}{\partial t} + \operatorname{div}(f(u)) = 0 \text{ for } (x, t) \in \mathbb{R}^d \times \mathbb{R}_+. \quad (13)$$

with

$$f(u) = (f_1(u), \dots, f_d(u)) \in C^1(\mathbb{R}; \mathbb{R}^d), \quad \operatorname{div} f(u) = \sum_{j=1}^d \frac{\partial}{\partial x_j} f_j(u), \quad (14)$$

We supply the SCL (13) with initial condition

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}^d, \quad (15)$$

and an entropy admissibility condition, which we choose as the Kruřkov entropy condition or an equivalent version of it.

3.2 Entropy Solution

It is well-known that the deterministic Cauchy problem (13), (15) admits, for each $u_0 \in L^1(\mathbb{R}^d) \cap BV(\mathbb{R})$, a unique entropy solution u (see, e.g., [28, 58, 11]). For every $t > 0$, $u(\cdot, t) \in L^1(\mathbb{R}^d)$. We require the (nonlinear) data-to-solution map

$$S : u_0 \mapsto u(\cdot, t) = S(t) u_0, \quad t > 0 \quad (16)$$

in our subsequent development. To state its properties, we introduce some additional notation: for a Banach-space E with norm $\|\cdot\|_E$, and for $0 < T \leq +\infty$, we denote by $C([0, T]; E)$ the space of bounded and continuous functions from $[0, T]$ with values in E , and by $L^p(0, T; E)$, $1 \leq p \leq +\infty$, the space of strongly Bochner measurable functions from $(0, T)$ to E such that for $1 \leq p < +\infty$

$$\|v\|_{L^p(0, T; E)} = \left(\int_0^T \|v(t)\|_E^p dt \right)^{\frac{1}{p}}, \quad (17)$$

respectively, if $p = \infty$,

$$\|v\|_{L^\infty(0, T; E)} = \operatorname{ess\,sup}_{0 \leq t \leq T} \|v(t)\|_E \quad (18)$$

are finite. The following existence result is classical. (see, for example, [35, Thms. 2.13, 2.14, Thm. 4.3] or also [15, 28, 29, 37, 42],

Theorem 1. *Assume that in the SCL (14)-(15) holds $f \in C^1(\mathbb{R}; \mathbb{R}^d)$, and the initial data u_0 satisfies*

$$u_0 \in Z := L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \cap BV(\mathbb{R}^d). \quad (19)$$

Then there holds:

- 1) *The SCL (14) - (15) admits a unique entropy solution $u \in L^\infty(\mathbb{R}^d \times (0, T))$.*
- 2) *For every $t > 0$, the (nonlinear) data-to-solution map $S(t)$ given by*

$$u(\cdot, t) = S(t)u_0$$

satisfies

- i) *$S(t) : L^1(\mathbb{R}^d) \rightarrow L^1(\mathbb{R}^d)$ is a (non-expansive) Lipschitz map, i.e.,*

$$\|S(t)u_0 - S(t)v_0\|_{L^1(\mathbb{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbb{R}^d)}. \quad (20)$$

- ii) *$S(t)$ maps $(L^1 \cap BV)(\mathbb{R}^d)$ into $(L^1 \cap BV)(\mathbb{R}^d)$ and*

$$TV(S(t)u_0) \leq TV(u_0) \quad \forall u_0 \in (L^1 \cap BV)(\mathbb{R}^d). \quad (21)$$

- iii) *There hold the L^∞ and L^1 stability bounds*

$$\|S(t)u_0\|_{L^\infty(\mathbb{R}^d)} \leq \|u_0\|_{L^\infty(\mathbb{R}^d)}; \quad (22)$$

$$\|S(t)u_0\|_{L^1(\mathbb{R}^d)} \leq \|u_0\|_{L^1(\mathbb{R}^d)}. \quad (23)$$

- iv) *The mapping $S(t)$ is a uniformly continuous mapping from $L^1(\mathbb{R}^d)$ into $C([0, \infty); L^1(\mathbb{R}^d))$, and*

$$\|S(\cdot)u_0\|_{C([0, T]; L^1(\mathbb{R}^d))} = \max_{0 \leq t \leq T} \|S(t)u_0\|_{L^1(\mathbb{R}^d)} \leq \|u_0\|_{L^1(\mathbb{R}^d)}. \quad (24)$$

Hyperbolic conservation laws exhibit *finite propagation speed* of perturbations. As a consequence, compactly supported initial data gives rise to solutions which are compactly supported for all time, however, with time-dependent supports. We present one version of such a “domain of influence” result, for the SCL (14) - (15).

Proposition 2. *For the Cauchy problem (14)-(15), assume that $f \in C^1(\mathbb{R}; \mathbb{R}^d)$ and that u_0 satisfies (19). Assume moreover that the initial data $u_0 \in Z$ has compact support: there exists a finite, positive constant R such that*

$$\text{supp}(u_0) \subset [-R, R]^d. \quad (25)$$

Then, for every $0 < t < \infty$, the unique entropy solution u of the Cauchy problem (14)-(15) is compactly supported as well, and with $\bar{M} := \|u_0\|_{L^\infty(\mathbb{R}^d)} < \infty$, there holds

$$\text{supp}(u(t)) \subset [-(R+tB), R+tB]^d, \quad (26)$$

where

$$B := \|\partial_u f\|_{C^0([- \bar{M}, \bar{M}]; \mathbb{R}^d)}, \quad (27)$$

denotes a upper bound on the maximal propagation speed.

3.3 Random Entropy Solution

Based on Theorem 1, we will now formulate the SCL (14) - (15) for random initial data $u_0(\omega; \cdot)$, with deterministic flux. To this end, we denote $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. We assume given a Lipschitz continuous deterministic flux f and random initial data u_0 , which satisfies the

Assumption 1. *[Assumptions on the random input data]*

1. *The random initial data u_0 is an $L^1(\mathbb{R}^d)$ -valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. It is in particular a strongly Bochner measurable map*

$$u_0 : (\Omega, \mathcal{F}) \mapsto (L^1(\mathbb{R}^d), \mathcal{B}(L^1(\mathbb{R}^d))). \quad (28)$$

2. *The map u_0 is also strongly Bochner measurable from $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the space $Z = L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \cap BV(\mathbb{R}^d)$ (taking values in the separable Banach space $L^1(\mathbb{R}^d)$, the random initial data u_0 is in particular separably valued in Z) introduced in (19), so that*

$$\Omega \ni \omega \mapsto u_0(\omega; \cdot) \in Z = (L^1 \cap L^\infty \cap BV)(\mathbb{R}^d) \quad (29)$$

is strongly Bochner measurable; here, we equip the space Z in (19), (29) with the norm

$$\|u\|_Z := \|u\|_{L^1(\mathbb{R}^d)} + \|u\|_{L^\infty(\mathbb{R}^d)} + TV(u). \quad (30)$$

3. *There holds a uniform bound: for some constant $0 < \bar{M} < \infty$,*

$$\|u_0(\omega; \cdot)\|_{L^\infty(\Omega; Z)} \leq \bar{M} < \infty, \quad (31)$$

4. The random initial data u_0 satisfies the bounded support assumption (25) with probability one, i.e. there exists a constant $0 < R < \infty$ such that

$$\text{supp}(u_0(\omega, \cdot)) \subset [-R, R]^d \quad \text{with probability } 1. \quad (32)$$

5. The flux function f is bounded on the set of states: for \bar{M} as in (31), item 3., there holds, with $S = [-\bar{M}, \bar{M}]$, the bound (27).

Since $L^1(\mathbb{R}^d)$ and $C^1(\mathbb{R}^d; \mathbb{R}^d)$ are separable, we may impose on the mapping (28) k -th moment conditions

$$\|u_0\|_{L^k(\Omega; L^1(\mathbb{R}^d))} < \infty, \quad k \in \mathbb{N}, \quad (33)$$

where the Bochner spaces are defined in Section 2. We consider the *random scalar conservation law* (RSCL)

$$\begin{cases} \partial_t u(\omega; x, t) + \text{div}_x(f(\omega; u(\omega; x, t))) = 0, & t > 0, \\ u(\omega; x, 0) = u_0(\omega; x), \end{cases} \quad x \in \mathbb{R}^d. \quad (34)$$

Definition 2. [47] A random field $u : \Omega \ni \omega \rightarrow u(\omega; x, t)$, i.e., a measurable mapping from (Ω, \mathcal{F}) to $C([0, T]; L^1(\mathbb{R}^d))$, is a *random entropy solution* of the SCL (34) with random initial data u_0 satisfying (28) - (33) for some $k \geq 2$ and with a *spatially homogeneous* flux $f(u)$ if it satisfies the following,

(i.) Weak solution:

For \mathbb{P} -a.e $\omega \in \Omega$, $u(\omega; \cdot, \cdot)$ satisfies the following integral identity

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \left(u(\omega; x, t) \varphi_t(x, t) + \sum_{j=1}^d f_j(\omega; u(\omega; x, t)) \frac{\partial}{\partial x_j} \varphi(x, t) \right) dx dt \\ + \int_{\mathbb{R}^d} u_0(x, \omega) \varphi(x, 0) dx = 0, \end{aligned} \quad (35)$$

for all test functions $\varphi \in C_0^1(\mathbb{R}^d \times [0, T])$.

(ii.) Entropy condition: For any pair of (deterministic) entropy η and entropy flux $Q(\cdot)$ i.e., η, Q_j with $j = 1, 2, \dots, d$ are functions such that η is convex and such that $Q'_j(\cdot) = \eta' f'_j(\cdot)$ for all j , and u satisfies the following inequality

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \left(\eta(u(\omega; x, t)) \varphi_t(x, t) + \sum_{j=1}^d Q_j(u(\omega; x, t)) \frac{\partial}{\partial x_j} \varphi(x, t) \right) dx dt \\ + \int_{\mathbb{R}^d} \eta(u_0(\omega; x)) \varphi(x, 0) dx \geq 0, \end{aligned} \quad (36)$$

for all deterministic test functions $0 \leq \varphi \in C_0^1(\mathbb{R}^d \times [0, T])$, \mathbb{P} -a.s.

We remark that it suffices to assume that (36) holds for all Kruřkov entropy functions $\eta(u) = |u - k|$, where k is any constant, which we assume throughout what follows.

Theorem 2. *[[28, Chap. 2, Thms. 5.1, 5.2]] Consider the SCL (14) - (15) with spatially homogeneous, bounded flux $f \in C^1(\mathbb{R}; \mathbb{R}^d)$ with random initial data $u_0 : \Omega \rightarrow L^1(\mathbb{R}^d)$ satisfying Assumption 1 and the k -th moment condition (33) for some integer $k \geq 2$. Then there exists a random entropy solution $u : \Omega \rightarrow C([0, T]; L^1(\mathbb{R}^d))$ which is “pathwise” unique, i.e., for \mathbb{P} -a.e. $\omega \in \Omega$, described in terms of the deterministic, nonlinear mapping $S(t)$ from Theorem 1 such that*

$$u(\omega; \cdot, t) = S(t)u_0(\omega; \cdot), \quad t > 0, \quad \mathbb{P}\text{-a.e. } \omega \in \Omega. \quad (37)$$

Moreover, $u : \Omega \rightarrow C(0, T; L^1(\mathbb{R}^d))$ is \mathbb{P} -a.s. separably valued and strongly Bochner measurable.

For every $k \geq 1$, for every $0 \leq t \leq T < \infty$, and for \mathbb{P} -a.e. $\omega \in \Omega$ holds

$$\|u\|_{L^k(\Omega; C(0, T; L^1(\mathbb{R}^d)))} \leq \|u_0\|_{L^k(\Omega; L^1(\mathbb{R}^d))}, \quad (38)$$

$$\|S(t)u_0(\omega; \cdot)\|_{(L^1 \cap L^\infty)(\mathbb{R}^d)} \leq \|u_0(\omega; \cdot)\|_{(L^1 \cap L^\infty)(\mathbb{R}^d)} \quad (39)$$

$$TV(S(\omega; t)u_0(\omega; \cdot)) \leq TV(u_0(\omega; \cdot)). \quad (40)$$

There exists $\bar{M} < \infty$ such that

$$\|u_0\|_{L^\infty(\Omega; L^\infty(\mathbb{R}^d))} \leq \bar{M}. \quad (41)$$

and

$$\sup_{0 \leq t \leq T} \|u(\omega; \cdot, t)\|_{L^\infty(\mathbb{R}^d)} \leq \bar{R} \quad \mathbb{P}\text{-a.e. } \omega \in \Omega. \quad (42)$$

Theorem 2 ensures the existence of a unique random entropy solution $u(\omega; x, t)$ with finite k -th moments for bounded random flux, provided that $u_0 \in L^k(\Omega, \mathcal{F}, \mathbb{P}; Z)$.

The deterministic maximum principle (22) and (41) imply, in addition, that \mathbb{P} -a.e. realization of the random entropy solution u takes values in the state space $S = [-\bar{M}, \bar{M}]$, for a.e. $x \in \mathbb{R}^d$ and for all $t > 0$.

4 Monte Carlo and Multi-level Monte Carlo Finite Volume Methods

We present the Multilevel Monte Carlo Finite Volume Method (MLMC-FVM) for scalar conservation laws. We introduce it in several steps: first, we discuss MC sampling of random initial data, second, FV discretization of the samples on the single, fixed triangulation and, finally, its multi-level extension on a hierarchy of possibly unstructured grids.

4.1 Monte-Carlo Method

The MC method for the SCL with random data $u_0(\omega; x)$ as in (28) - (31) consists in sampling in the probability space. We also assume (33), i.e., the existence of k -th moments of u_0 for some $k \in \mathbb{N}$, to be specified. We analyze the error in computable numerical approximations of k -th order statistical moments of u . For $k = 1$ we obtain the expected value of the solution random field i.e, $\mathcal{M}^1(u) = \mathbb{E}[u]$. The *MC approximation of $\mathbb{E}[u]$* is defined as the usual statistical sample average: given M independent, identically distributed (“iid” for short) draws of the initial data, \widehat{u}_0^i , $i = 1, \dots, M$, the MC estimate of $\mathbb{E}[u(\cdot; \cdot, t)]$ at time $t > 0$ is the sample average

$$E_M[u(\cdot, t)] := \frac{1}{M} \sum_{i=1}^M \widehat{u}^i(\cdot, t). \quad (43)$$

Here, $\widehat{u}^i(\cdot, t)$ denotes the M unique entropy solutions of the M Cauchy Problems (14) - (15) with iid initial data \widehat{u}_0^i . We observe that by

$$\widehat{u}^i(\cdot, t) = \widehat{S}(t) \widehat{u}_0^i \quad (44)$$

we have from (21) - (23) for M MC samples and for any $0 < t < \infty$, for every $1 \leq p \leq \infty$, using the triangle inequality,

$$\|E_M[u(\omega; \cdot, t)]\|_{L^p(\mathbb{R}^d)} = \left\| \frac{1}{M} \sum_{i=1}^M \widehat{S}(t) \widehat{u}_0^i(\cdot; \omega) \right\|_{L^p(\mathbb{R}^d)} \leq \frac{1}{M} \sum_{i=1}^M \|\widehat{u}_0^i(\omega; \cdot)\|_{L^p(\mathbb{R}^d)}. \quad (45)$$

Using the i.i.d. property of the samples $\{\widehat{u}_0^i\}_{i=1}^M$ of the random initial data u_0 and the linearity of the expectation $\mathbb{E}[\cdot]$, we obtain for any $1 \leq p \leq \infty$ from the *assumed strong measurability* of u_0 in the Banach space Z defined in (29), the bound

$$\mathbb{E} \left[\|E_M[u(\cdot; \cdot, t)]\|_{L^p(\mathbb{R}^d)} \right] \leq \mathbb{E} \left[\|u_0\|_{L^p(\mathbb{R}^d)} \right] = \|u_0\|_{L^1(\Omega; L^p(\mathbb{R}^d))} < \infty. \quad (46)$$

As $M \rightarrow \infty$, the MC estimates (43) converge in $L^2(\Omega; C([0, T]; L^p(\mathbb{R}^d)))$ and the following convergence rate bound holds.

Theorem 3. *Assume that in the SCL (14) - (15) the random initial data u_0 satisfies Assumption 1, items 1.-5. In particular, u_0 is with probability one compactly supported in space, i.e, there exists a compact domain $D \subset \mathbb{R}^d$ such that $\text{supp}(u_0(\omega)) \subset D$ for almost every $\omega \in \Omega$.*

Assume, moreover, that the random initial data u_0 is strongly Bochner measurable taking values in the space $Z = L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \cap BV(\mathbb{R}^d)$ (cp. (29), (30)), and satisfies

$$u_0 \in L^2(\Omega; L^1(\mathbb{R}^d)) \cap L^2(\Omega; L^\infty(\mathbb{R}^d)). \quad (47)$$

Assume further that (29), (31) hold.

Then for every $0 < t < \infty$ holds the apriori bound

$$\|u(t)\|_{L^2(\Omega;L^2(\mathbb{R}^d))}^2 \leq \|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \|u_0\|_{L^2(\Omega;L^\infty(\mathbb{R}^d))} \quad (48)$$

The MC estimates $E_M[u(\cdot, t)]$ in (43) converge, as $M \rightarrow \infty$, to $\mathcal{M}^1(u(\cdot, t)) = \mathbb{E}[u(\cdot, t)]$. For $M \in \mathbb{N}$, and for every fixed $0 < t < \infty$, there holds the error bound

$$\|\mathbb{E}[u(\cdot, t)] - E_M[u(\cdot, t)]\|_{L^2(\Omega;L^2(\mathbb{R}^d))}^2 \leq M^{-1} \|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \|u_0\|_{L^2(\Omega;L^\infty(\mathbb{R}^d))}. \quad (49)$$

Under Assumption 1, item 4., we also have for every $0 < t < T$

$$\|\mathbb{E}[u(\cdot, t)] - E_M[u(\cdot, t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))}^2 \leq C(T)M^{-1} \|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \|u_0\|_{L^2(\Omega;L^\infty(\mathbb{R}^d))}. \quad (50)$$

In (50), $C(T)$ is a time dependent constant that also depends on the bounded domain $[-R, R]^d$ on which the random initial data is supported with probability 1.

Proof. Under the assumptions of the theorem, by Theorem 2 there exists a unique random entropy solution u .

For any $v \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ holds $\|v\|_{L^2(\mathbb{R}^d)}^2 \leq \|v\|_{L^1(\mathbb{R}^d)} \|v\|_{L^\infty(\mathbb{R}^d)}$. For every fixed $0 < t < \infty$ we have from (22), (23),

$$\begin{aligned} \int_{\Omega} \|S(t)u_0\|_{L^2(\mathbb{R}^d)}^2 &= \int_{\Omega} \|u(\omega; t)\|_{L^2(\mathbb{R}^d)}^2 d\mathbb{P}(\omega) \\ &\leq \int_{\Omega} \|u_0(\omega)\|_{L^1(\mathbb{R}^d)} \|u_0(\omega)\|_{L^\infty(\mathbb{R}^d)} d\mathbb{P}(\omega) \\ &\leq \|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \|u_0\|_{L^2(\Omega;L^\infty(\mathbb{R}^d))}. \end{aligned}$$

Therefore, for every $0 < T < \infty$,

$$\begin{aligned} \|u\|_{L^2(\Omega;C(0,T;L^2(\mathbb{R}^d)))}^2 &= \int_{\Omega} \sup_{0 < t < T} \|u(\omega; t)\|_{L^2(\mathbb{R}^d)}^2 d\mathbb{P}(\omega) \\ &\leq \sup_{0 < t < T} \|S(t)u_0\|_{L^2(\Omega;L^2(\mathbb{R}^d))}^2 \\ &\leq \|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \|u_0\|_{L^2(\Omega;L^\infty(\mathbb{R}^d))}, \end{aligned} \quad (51)$$

which is finite by assumption (47). From this bound follows the MC error bound (49) by referring to the general MC error bound in Corollary 1, with the observation that Hilbert spaces have type $p = 2$, and type constant $2C_t = 1$ or directly from (12).

We show the second part: note that the space $L^1(\mathbb{R}^d)$ is of type 1. From Corollary 1 we can not expect a MC convergence rate bound in $L^1(\mathbb{R}^d)$ without extra assumptions.

Suppose therefore now that Assumption 1, item 4., holds, ie. all realizations of u_0 have compact support in a common set $[-R, R]^d$. Then the bound (27) of the flux f in $C^1(S; \mathbb{R}^d)$ and the finite propagation property, Prop. 2, imply that for every $0 < t < \infty$, and \mathbb{P} -a.s., that the random entropy solution is likewise compactly supported: from (26) it follows that there holds, for every $t > 0$,

$$\text{supp}(u(\omega; t)) \subset [-(R+tB), R+tB]^d \quad \text{with probability 1}. \quad (52)$$

Then, for \mathbb{P} -a.e. ω ,

$$\|u(\omega; t)\|_{L^1(\mathbb{R}^d)} \leq C(t, B, R) \|u(\omega; t)\|_{L^2(\mathbb{R}^d)}.$$

Squaring both sides and taking expectations, we find

$$\|u(t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))}^2 \leq C(t, B, R)^2 \|u(\omega; t)\|_{L^2(\Omega;L^2(\mathbb{R}^d))}^2.$$

Using (48) and reasoning as before, we arrive at (50). \square

Remark 2. The bound (51) can be generalized to k -point correlation functions $\mathcal{M}^{(k)}u = \mathbb{E}(u(\cdot, t, x_1) \dots u(\cdot, t, x_k))$, $x_1, \dots, x_k \in \mathbb{R}^d$ with $k > 1$ of the random entropy solution: from Jensen's inequality and Fubini's theorem,

$$\begin{aligned} \|\mathcal{M}^{(k)}u(t)\|_{L^2(\mathbb{R}^{kd})}^2 &\leq \int_{\Omega} \int_{x_1} \dots \int_{x_k} |u(\cdot, x_1, t) \dots u(\cdot, x_k, t)|^2 dx_k \dots dx_1 d\mathbb{P}(\omega) \\ &= \int_{\Omega} \|u(\cdot, t, \cdot)\|_{L^2(\mathbb{R}^d)}^{2k} d\mathbb{P}(\omega) \\ &= \int_{\Omega} \|S(t)u_0(\cdot)\|_{L^2(\mathbb{R}^d)}^{2k} d\mathbb{P}(\omega) \\ &\leq \int_{\Omega} \|S(t)u_0(\cdot)\|_{L^1(\mathbb{R}^d)}^k \|S(t)u_0(\cdot)\|_{L^\infty(\mathbb{R}^d)}^k d\mathbb{P}(\omega) \\ &\leq \int_{\Omega} \|u_0(\cdot)\|_{L^1(\mathbb{R}^d)}^k \|u_0(\cdot)\|_{L^\infty(\mathbb{R}^d)}^k d\mathbb{P}(\omega) \\ &\leq \|u_0\|_{L^{2k}(\Omega;L^1(\mathbb{R}^d))}^k \|u_0\|_{L^{2k}(\Omega;L^\infty(\mathbb{R}^d))}^k. \end{aligned}$$

From Theorem 3 we see that $L^1(\mathbb{R}^d)$ MC convergence rate bounds can be obtained despite L^1 being a Banach space of type 1; however, as already observed in [13] and the references there, this is only possible by an intermediate bound of samples and, for multilevel MC, for error bounds on FV discretizations in Banach spaces of type $1 < q \leq 2$, as introduced in Def. 1. As observed in [46], Theorem 4.1, the assumption of compactly supported initial data, satisfying (25), and bounded flux (27) which imply (26) at positive time $t > 0$ does afford intermediate $L_\omega^2 L_x^2$ bounds which in turn allow MC convergence rate bounds.

The properties (20) - (24) also hold for FV discretizations. Accordingly, we aim at analogous results for MC FV methods for the random SCL. We next introduce FV methods; rather than presenting a particular scheme, we state several properties required in the ensuing error analysis which are satisfied by several popular FV methods.

4.2 Finite Volume Method (FVM)

So far, the MC method was prescribed under the assumption that ‘‘pathwise’’ entropy solutions $\hat{u}^i(\omega; x, t) = S(\omega; t)\hat{u}_0^i(\omega; x)$ for the Cauchy problem (14) - (15) iid initial data samples $\hat{u}_0^i = u_0(\omega; x)$ are available exactly. In practice, numerical approximations of $S(t)\hat{u}_0^i$ must be computed and the corresponding discretization errors taken into account.

To analyze MC-FVM approximations we impose sufficient conditions on the FVM to afford the Kuznetsov type error bounds for first order FVM for the *deterministic*

SCL (14)-(15); these will be required for the convergence rate bounds of the MLMC FVM as considered in [46, 47] and also for parametric collocation FVM as in [46, Sec.5] in the subsequent chapters. We review the generic first order, explicit FV schemes considered here, as for example in [15, 42, 37].

Denote the time step by $\Delta t > 0$ and a triangulation \mathcal{T} of the spatial domain $D \subset \mathbb{R}^d$ of interest. We assume that \mathcal{T} is a set of open, convex polyhedra $K \subset \mathbb{R}^d$ with plane faces such that standard conditions on shape regularity hold: if $K \in \mathcal{T}$ denotes a generic volume, we define

$$\rho_K = \rho(K) = \max\{\text{diam}(B_r) : B_r \subset \bar{K}\} \quad (53)$$

i.e., the maximum inradius in volume \bar{K} for $K \in \mathcal{T}$ and define, in addition, for a generic mesh \mathcal{T} , the *shape regularity constants* (where $\Delta x_K := \text{diam}(K)$)

$$\kappa(\mathcal{T}) := \sup\{\Delta x_K / \rho(K) : K \in \mathcal{T}\}, \quad \mathcal{T} \in \mathfrak{M}. \quad (54)$$

The meshwidth of triangulation \mathcal{T} is

$$\Delta x(\mathcal{T}) = \sup\{\text{diam}(K) : K \in \mathcal{T}\}. \quad (55)$$

For any volume $K \in \mathcal{T}$, we define the *set $\mathcal{N}(K)$ of neighboring volumes*

$$\mathcal{N}(K) := \{K' \in \mathcal{T} : K' \neq K \wedge \text{meas}_{d-1}(\bar{K} \cap \bar{K}') > 0\}. \quad (56)$$

We assume that the triangulation \mathcal{T} are regular in the sense that the support size of the FV “stencil” at any element $K \in \mathcal{T}$ is uniformly bounded i.e.,

$$\sigma(\mathcal{T}) := \sup_{K \in \mathcal{T}} \#\mathcal{N}(K) \leq B < \infty \quad (57)$$

with some bound B which is independent of the particular partition \mathcal{T} . The global CFL constant is defined by

$$\lambda := \Delta t / \Delta x(\mathcal{T}). \quad (58)$$

for constant time step Δt ; we also set $t_n = n\Delta t$. It is determined by a standard CFL condition (see e.g. [28]) based on the maximum wave speed given by the flux bound (27), see Proposition 2).

We discretize (14)-(15) by an explicit, first order FV scheme on \mathcal{T} :

$$v_K^{n+1} = H(\{v_{K'}^n : K' \in \mathcal{N}(K) \cup K\}), \quad K \in \mathcal{T} \quad (59)$$

where $H : \mathbb{R}^{(2k+1)^d} \rightarrow \mathbb{R}$, with k denoting the size of the stencil of the FV scheme, is continuous and where v_K^n denotes an approximation to the cell average of u at time $t_n = n\Delta t$.

In our subsequent developments, we write the FVM in operator form. To this end, we introduce the operator $H_{\mathcal{T}}(v)$ which maps a sequence $\underline{v} = (v_K)_{K \in \mathcal{T}}$ into $H_{\mathcal{T}}((v_K)_{K \in \mathcal{T}})$. The time explicit FVM (59) takes the abstract form

$$\underline{v}^{n+1} = H_{\mathcal{T}}(\underline{v}^n), \quad n = 0, 1, 2, \dots \quad (60)$$

For the ensuing convergence analysis, we shall assume and use several properties of the FV scheme (60); these properties are satisfied by many commonly used FVM of the form (60), on regular or irregular meshes \mathcal{T} in \mathbb{R}^d .

To state the assumptions, we introduce further notation: for any initial data $u_0(x) \in L^1(\mathbb{R}^d)$, we define the FVM approximation at time $t = 0$, $(v_K^0)_{K \in \mathcal{T}}$ by the *cell averages*

$$v_K^0 = \frac{1}{|K|} \int_K u_0(x) dx, \quad \text{where } K \in \mathcal{T}. \quad (61)$$

Interpreting the vector $\underline{v} = (v_K)_{K \in \mathcal{T}} \in \mathbb{R}^{\#\mathcal{T}}$ as cell averages, we associate with \underline{v} the piecewise constant function $v_{\mathcal{T}}(x, t)$ defined a.e. in $\mathbb{R}^d \times (0, \infty)$ by

$$v_{\mathcal{T}}(x, t)|_K := v_K^n, \quad K \in \mathcal{T}, \quad t \in [t_n, t_{n+1}). \quad (62)$$

We denote space of all piecewise constant functions on \mathcal{T} (i.e., the “simple” or “step” functions on \mathcal{T}) by $S(\mathcal{T})$. Given any $v_{\mathcal{T}} \in S(\mathcal{T})$, we define the (mesh-dependent) norms:

$$\|\underline{v}\|_{L^1(\mathcal{T})} = \sum_{K \in \mathcal{T}} |K| |v_K| = \|v_{\mathcal{T}}\|_{L^1(\mathbb{R}^d)}, \quad \|\underline{v}\|_{L^\infty(\mathcal{T})} = \sup_{K \in \mathcal{T}} |v_K| = \|v_{\mathcal{T}}\|_{L^\infty(\mathbb{R}^d)}.$$

As in [47, 46], we consider FVM schemes in the MC-FVM algorithms which satisfy the following standard assumptions which are analogous to (22), (23) and (24).

Assumption 2. *The abstract FV scheme (60) satisfies*

1. Stability: $\forall t \geq 0$

$$\|v_{\mathcal{T}}(\cdot, t)\|_{L^\infty(\mathbb{R}^d)} \leq \|v_{\mathcal{T}}(\cdot, 0)\|_{L^\infty(\mathbb{R}^d)}, \quad (63)$$

$$\|v_{\mathcal{T}}(\cdot, t)\|_{L^1(\mathbb{R}^d)} \leq \|v_{\mathcal{T}}(\cdot, 0)\|_{L^1(\mathbb{R}^d)}, \quad (64)$$

$$TV(v_{\mathcal{T}}(\cdot, t)) \leq TV(v_{\mathcal{T}}(\cdot, 0)), \quad (65)$$

2. Lipschitz continuity: *For any two sequences $\underline{v} = (v_K)_{K \in \mathcal{T}}$, $\underline{w} = (w_K)_{K \in \mathcal{T}}$ we have*

$$\|H_{\mathcal{T}}(\underline{v}) - H_{\mathcal{T}}(\underline{w})\|_{L^1(\mathcal{T})} \leq \|\underline{v} - \underline{w}\|_{L^1(\mathcal{T})} \quad (66)$$

or, equivalently,

$$\|H_{\mathcal{T}}(v_{\mathcal{T}}) - H_{\mathcal{T}}(w_{\mathcal{T}})\|_{L^1(\mathbb{R}^d)} \leq \|v_{\mathcal{T}} - w_{\mathcal{T}}\|_{L^1(\mathbb{R}^d)}. \quad (67)$$

3. Convergence: *If in the CFL bound (58) the CFL constant $\lambda = \Delta t / \Delta x(\mathcal{T})$ is kept constant as $\Delta x(\mathcal{T}) \rightarrow 0$, the approximate solution $v_{\mathcal{T}}(x, t)$ generated by (59) - (62) converges to the unique entropy solution u of the scalar conservation laws (14) - (15) at $L^1(\mathbb{R}^d)$ -rate $0 < s \leq 1$: there exists $C > 0$ independent of Δx such that, as $\Delta x \rightarrow 0$, for every \bar{t} and for $(\Delta t)^s \leq \bar{t} \leq T$, it holds*

$$\|u(\cdot, \bar{t}) - v_{\mathcal{T}}(\cdot, \bar{t})\|_{L^1(\mathbb{R}^d)} \leq \|u_0 - v_{\mathcal{T}}^0\|_{L^1(\mathbb{R}^d)} + C\bar{t}TV(u_0)\Delta x^s. \quad (68)$$

Let us mention that (63), (64) and (65) hold in particular for monotone schemes on Cartesian meshes, see [28, 37]. The classical error analysis of Kuznetsov, see e.g. [15], imply the convergence rate $s = 1/2$ in (68). In case of monotone schemes on general FV meshes, one might lose the bound on the total variation of the approximations, and the convergence rate, i.e., the rate s in (68) is correspondingly reduced, see [7].

The error bound (68) contains an initial data approximation error $\|u_0 - v_{\mathcal{T}}^0\|_{L^1(\mathbb{R}^d)}$. This error vanishes for step function initial data on \mathcal{T} (as, e.g., in the solution of Riemann problems). More generally, this error can be bounded by Δx^s provided that u_0 has appropriate regularity: under Assumption 2, for $u_0 \in BV(\mathbb{R}^d)$ and for the cell-average projection $v_{\mathcal{T}}^0$ in (61), we obtain

$$\|u_0 - v_{\mathcal{T}}^0\|_{L^1(\mathbb{R}^d)} \leq C(\kappa, \sigma)\Delta xTV(u_0) \leq C(\kappa, \sigma)\Delta x^sTV(u_0), \quad (69)$$

as $s \leq 1$ in (68). Here, the constant $C(\kappa, \sigma)$ depends on the stencil size constant σ and the shape regularity constant κ in (57), and (54), respectively.

Explicit FV schemes (59), (60) subject to the CFL stability condition (58) exhibit a discrete finite domain of dependence result analogous to Proposition 2.

Proposition 3. *[Discrete finite dependence domain]*

Under Assumption 2 and the assumptions of Proposition 2, in particular under the compact support assumption (25) on the random initial data u_0 , for the explicit FV scheme (53), (61) there holds:

1. the projection of the initial data on triangulation \mathcal{T} , $v_{\mathcal{T}}^0$, defined in (61), (62), has compact support independently of \mathcal{T} : there exists $c > 0$ such that, for all $0 < h(\mathcal{T}) \leq 1$, and with probability 1, holds

$$\text{supp}(v_{\mathcal{T}}^0) \subset [-(1+c)R, (1+c)R]^d. \quad (70)$$

2. the discrete solutions satisfy a dependence domain result: with probability 1 and with the constant $c > 1$ from (70) for every $t > 0$ holds

$$\text{supp}(v_{\mathcal{T}}(\cdot, t)) \subset [-(1+c)(R+tB), (1+c)(R+tB)]^d, \quad (71)$$

where B denotes the bound (27) on the flux, and where $c > 1$ is as in (70).

We refer to [42, Chapter 3.6] for a detailed discussion.

Let us finally mention that the work for the realization of scheme (59) - (62) on a bounded domain $D \subset \mathbb{R}^d$ scales as (using the CFL stability condition (58), i.e., $\Delta t/\Delta x(\mathcal{T}) \leq \lambda = \text{const.}$)

$$\text{Work}_{\mathcal{T}} = \mathcal{O}\left(\Delta t^{-1}\Delta x^{-d}\right) = \mathcal{O}\left(\Delta x^{-(d+1)}\right), \quad (72)$$

with the constant implied in $\mathcal{O}(\cdot)$ depending on on the support domain D of the solution.

4.3 MC-FVM

In the Monte Carlo Finite Volume Methods (MC-FVMs), we combine MC sampling of the random initial data with the FVM (60). In the convergence analysis of these schemes, we shall require the application of the FVM (60) to random initial data $u_0 \in L^\infty(\Omega; (L^1 \cap L^\infty \cap BV)(\mathbb{R}^d))$: given a draw $u_0(x; \omega)$ of u_0 , the FVM (60) - (62) produces a family $v_{\mathcal{T}}(x, t; \omega)$ of random step functions on \mathcal{T} .

Proposition 4. *Consider the FVM (60) - (62) for the approximation of the entropy solution corresponding to a draw $u_0(\omega; x)$ of the random initial data, satisfying Assumption 1.*

Then, if the FVM satisfies Assumption 2, and provided that $u_0 \in L^k(\Omega; Z)$, the random grid functions $\Omega \ni \omega \mapsto v_{\mathcal{T}}(\omega; x, t)$ defined by (58) - (62) satisfy, for every $0 < \bar{t} < \infty$, $0 < \Delta x < 1$, and every $k \in \mathbb{N} \cup \{\infty\}$, the stability bounds

$$\|v_{\mathcal{T}}(\cdot; \cdot, \bar{t})\|_{L^k(\Omega; L^\infty(\mathbb{R}^d))} \leq \|u_0\|_{L^k(\Omega; L^\infty(\mathbb{R}^d))}, \quad (73)$$

$$\|v_{\mathcal{T}}(\cdot; \cdot, \bar{t})\|_{L^k(\Omega; L^1(\mathbb{R}^d))} \leq \|u_0\|_{L^k(\Omega; L^1(\mathbb{R}^d))}. \quad (74)$$

We also have error bounds

$$\begin{aligned} & \|u(\cdot; \cdot, \bar{t}) - v_{\mathcal{T}}(\cdot; \cdot, \bar{t})\|_{L^k(\Omega; L^1(\mathbb{R}^d))} \\ & \leq \|u_0(\cdot; \cdot) - v_{\mathcal{T}}^0(\cdot; \cdot)\|_{L^k(\Omega; L^1(\mathbb{R}^d))} + C\bar{t}\Delta x^s \|TV(u_0(\cdot; \cdot))\|_{L^k(\Omega)}. \end{aligned} \quad (75)$$

Remark 3. 1. In order for $\|TV(u_0(\cdot; \cdot))\|_{L^k(\Omega)}$ in (75) to be meaningful, a sufficient condition is that $u_0 : \Omega \rightarrow BV(\mathbb{R}^d)$ be strongly measurable, which we assumed in (31).

2. The initial data approximation error term $\|u_0(\cdot; \cdot) - v_{\mathcal{T}}^0(\cdot; \cdot)\|_{L^k(\Omega; L^1(\mathbb{R}^d))}$ in (75) can be bounded as in (69) provided that the random initial data u_0 has sufficient regularity: if $u_0 : \Omega \rightarrow Z$ is strongly measurable, and if $u_0 \in L^k(\Omega; Z)$, then (69) implies

$$\|u_0 - v_{\mathcal{T}}^0\|_{L^k(\Omega; L^1(\mathbb{R}^d))} \leq C\Delta x. \quad (76)$$

This approximation error bound holds without assumption of bounded support on u_0 .

4.3.1 Definition of the MC-FVM Scheme

We consider once more the SCL (14) - (15) with random data u_0 and with flux f satisfying (28) - (33). We assume the moment condition $u_0 \in L^k(\Omega; Z)$ for sufficiently large $k \in \mathbb{N}$. The MC-FVM scheme for the MC estimation of the mean (ie., the ensemble average) of the random entropy solution is as follows.

Definition 3. (MC-FVM Scheme) Generate a sample of $M \in \mathbb{N}$ i.i.d. realizations $\{\hat{u}_0^i\}_{i=1}^M$ of initial data, approximated on the triangulation \mathcal{T} by cell average projections (61).

$$\widehat{u}^i(\cdot, t) = S(t)\widehat{u}_0^i(\cdot), \quad i = 1, \dots, M. \quad (77)$$

Let $H_{\mathcal{T}}(\cdot)$ be a FVM scheme (59) - (62) satisfying Assumption 2. Then the MC-FVM approximations of $\mathcal{M}^k(u(\cdot, t))$ are defined as statistical estimates from the ensemble

$$\{\widehat{v}_{\mathcal{T}}^i(\cdot, t)\}_{i=1}^M \quad (78)$$

obtained by (60) from the FV approximations $\widehat{v}_{\mathcal{T}}^i(\cdot, 0)$ of the M i.i.d initial data samples $\{\widehat{u}_0^i(x)\}_{i=1}^M$ by (61): specifically, the first moment of the random solution $u(\omega; \cdot, t)$ at time $t > 0$, is approximated by the sample average of FV solutions,

$$\mathcal{M}^1(u(\cdot, t)) \approx E_M[v_{\mathcal{T}}(\cdot, t)] := \frac{1}{M} \sum_{i=1}^M \widehat{v}_{\mathcal{T}}^i(\cdot, t). \quad (79)$$

4.3.2 Convergence Rates for MC-FVM

We next address the convergence of $E_M[v_{\mathcal{T}}]$ to the mean $\mathbb{E}(u)$.

Theorem 4. *Assume that all realizations of the random initial data u_0 are supported on one common, bounded domain $[-R, R]^d \subset \mathbb{R}^d$ for some $0 < R < \infty$ and satisfy (28) - (32). Assume further given a FVM (59) - (62) such that (58) holds and such that Assumption 2 is satisfied; in particular, assume that the deterministic FVM scheme converges at rate $s > 0$ in $C([0, T]; L^1(\mathbb{R}^d))$ for every $0 < T < \infty$, i.e. (68) holds.*

Then, the MC estimate $E_M[v_{\mathcal{T}}(\cdot, t)]$ defined in (79) satisfies, for every M , the error bound

$$\begin{aligned} & \|\mathbb{E}[u(\cdot, t)] - E_M[v_{\mathcal{T}}(\omega; \cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ & \leq C(D, T) \left[M^{-\frac{1}{2}} \|u_0\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \right. \\ & \quad \left. + \|u_0 - v_{\mathcal{T}}^0\|_{L^\infty(\Omega; L^1(\mathbb{R}^d))} + t \Delta x^s \|TV(u_0(\omega; \cdot))\|_{L^\infty(\Omega)} \right] \end{aligned} \quad (80)$$

where $C > 0$ depends on the final time T and the domain D , in which the initial data is supported \mathbb{P} -a.s. but is independent of M and of Δx as $M \rightarrow \infty$ and as $\lambda \Delta x = \Delta t \downarrow 0$.

Proof. The proof of the above theorem proceeds along the lines of the proof of [47, Thm. 4.6]. However, we point out that there was an error in the argument of the proof of [47, Theorem 4.6] due to the incorrect derivation of a direct MC sampling convergence rate in the type one Banach space $L^1(\mathbb{R}^d)$. On the other hand and as mentioned in the previous section, we may use the local support assumptions on the initial data, and the finite speed of propagation implied by hyperbolicity, to obtain FV convergence rate bounds in the type 2 space $L^2(\mathbb{R}^d)$ from which follows the claimed convergence rate.

For fixed $t > 0$, we have

$$\begin{aligned} \|\mathbb{E}[u(\cdot, t)] - E_M[v_{\mathcal{T}}(\cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} &\leq \|\mathbb{E}[u(\cdot, t)] - E_M[u(\cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ &\quad + \|E_M[u(\cdot, t)] - E_M[v_{\mathcal{T}}(\cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ &=: \text{I} + \text{II} \quad . \end{aligned}$$

Term I is a MC error which can be bounded by (50).

Term II is essentially a discretization error bound. By the (pathwise) FV error bounds (68) and by (21) - (24) and Assumption 2 by the triangle inequality that

$$\begin{aligned} \text{II} &= \|E_M[u(\cdot, t; \omega) - v_{\mathcal{T}}(\cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ &\leq \frac{1}{M} \sum_{j=1}^M \|\hat{u}^j(\cdot, t; \omega) - \hat{v}_{\mathcal{T}}^j(\cdot, t; \omega)\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ &\leq \|u(\cdot, t; \omega) - v_{\mathcal{T}}(\cdot, t; \omega)\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ &\leq C \{ \|u_0 - v_{\mathcal{T}}^0\|_{L^2(\Omega; L^1(\mathbb{R}^d))} + t \Delta x^s \|TV(u_0(\cdot, w))\|_{L^2(\Omega)} \} . \end{aligned}$$

□

The initial data approximation error $\|u_0 - v_{\mathcal{T}}^0\|_{L^2(\Omega; L^1(\mathbb{R}^d))}$ can be bounded by Δx as indicated in Remark 3, item 2.

4.3.3 Work estimates

To calculate the error versus computational work, we estimate the asymptotic complexity of computing the estimators along the lines of [46]. In doing this, we *assume that the computational domain $D \subset \mathbb{R}^d$ is bounded and suitable boundary conditions are specified on ∂D* . In this *bounded, computational domain D* , the work for one time step (59), (60) is of order $\mathcal{O}(\Delta x^{-d})$ (with $\mathcal{O}(\cdot)$ depending on the size of the domain and on the size of stencil employed in the FV scheme) we find from the CFL condition (58) that the total computational work to obtain $\{v_{\mathcal{T}}(\cdot, t)\}_{0 < t \leq T}$ in D is by (72)

$$\text{Work}(\mathcal{T}) = \mathcal{O}(\Delta x^{-d-1}), \quad \text{as } \lambda \Delta x = \Delta t \downarrow 0. \quad (81)$$

The work for the computation of the MC estimate $E_M[v_{\mathcal{T}}(\cdot, t)]$ is assumed to scale as

$$\text{Work}(M, \mathcal{T}) = \mathcal{O}(M \Delta x^{-d-1}), \quad \text{as } \Delta t = \lambda \Delta x \downarrow 0. \quad (82)$$

The bound (80) allows to infer convergence order estimates in terms of work. To derive these, we choose $M^{-1/2} \sim \Delta x^s$ in (80). Setting the implied constant equal to one results in $M = \Delta t^{-2s}$. Inserting in (82) yields

$$\text{Work}(\mathcal{T}) = \mathcal{O}(\Delta t^{-2s} \Delta x^{-(d+1)}) \stackrel{(58)}{=} \mathcal{O}(\Delta x^{-(d+1)-2s}) \quad (83)$$

so that we obtain from (80)

$$\|\mathbb{E}[u(\cdot, t)] - E_M[v_{\mathcal{T}}(\cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \leq C\Delta t^s \leq C(\text{Work}(\mathcal{T}))^{-s/(d+1+2s)}. \quad (84)$$

Regarding the convergence rate in the estimate (84), for the *deterministic FV scheme* holds

$$\text{Work}(\mathcal{T}) = \mathcal{O}\left(\Delta t^{-1} \Delta x^{-d}\right) \stackrel{(58)}{=} \mathcal{O}\left(\Delta x^{-(d+1)}\right).$$

The bound on the FV error, (68), becomes when written in terms of work, equal to

$$\begin{aligned} & \|u(\cdot, \bar{t}) - v_{\mathcal{T}}(\cdot, \bar{t})\|_{L^1(\mathbb{R}^d)} \\ & \leq \|u_0 - v_{\mathcal{T}}^0\|_{L^1(\mathbb{R}^d)} + C\bar{t}TV(u_0) (\text{Work}(\mathcal{T}))^{-s/(d+1)}. \end{aligned} \quad (85)$$

Ignoring initial data approximation errors, which are negligible in comparison to the computational work for the time marching, the exponent $-s/(d+1)$ for the deterministic FVM as compared to $-s/(d+1+2s)$ for the MC-FVM. For low order FV schemes (ie., for small values of the convergence rate s) and in space dimensions $d = 2, 3$, we observe a considerably reduced rate of convergence of the MC-FVM. For high order FV schemes, we recover in (84) the MC rate $1/2$ of the error in terms of work.

4.4 Multilevel MC-FVM

Next, we present and analyze a scheme that allows us to achieve a better accuracy versus work bound for the random initial data u_0 , compared to the standard MC-FVM error bound (84). The Multilevel Monte Carlo Finite Volume (MLMC-FVM) scheme is based on MC sampling with level dependent sample sizes M_ℓ on different levels ℓ of resolution of the FVM. Throughout, we assume the explicit FV scheme satisfies Assumption 2 and the CFL stability condition (58). To define the MLMC-FVM, we start by reviewing notation as used in [46].

4.4.1 Notation

The MLMC-FVM is defined as a *multilevel discretization* in x and t with level dependent numbers M_ℓ of samples. To this end, we assume we are given a family $\{\mathcal{T}_\ell\}_{\ell=0}^\infty$ of *nested triangulations* of \mathbb{R}^d such that the mesh width

$$\Delta x_\ell = \Delta x(\mathcal{T}_\ell) = \sup\{\text{diam}(K) : K \in \mathcal{T}_\ell\} = \mathcal{O}\left(2^{-\ell} \Delta x_0\right), \quad \ell \in \mathbb{N}_0, \quad (86)$$

where K denotes a generic FV cell $K \in \mathcal{T}$. We also assume the family $\mathfrak{M} = \{\mathcal{T}_\ell\}_{\ell=0}^\infty$ of meshes to be shape regular; if $K \in \mathcal{T}_\ell$ denotes a generic cell, we recall, for a generic mesh $\mathcal{T} \in \mathfrak{M}$, the *shape regularity constants* $\kappa(\mathcal{T})$ defined in (54). We say that *the family \mathfrak{M} of meshes is κ -shape regular*, if there exists a constant $\kappa(\mathfrak{M}) < \infty$ such that with ρ_K denoting the diameter of the largest ball inscribed into K

$$\kappa(\mathfrak{M}) = \sup_{\mathcal{T} \in \mathfrak{M}} \kappa(\mathcal{T}) = \sup_{\mathcal{T} \in \mathfrak{M}} \sup_{K \in \mathcal{T}} \frac{\text{diam}(K)}{\rho_K}. \quad (87)$$

For a mesh hierarchy $\mathfrak{M} = \{\mathcal{T}_\ell\}_{\ell=0}^\infty$, we denote

$$\Delta x_\ell := \Delta x(\mathcal{T}_\ell), \quad \mathcal{T}_\ell \in \mathfrak{M}, \quad \ell = 0, 1, \dots \quad (88)$$

4.4.2 MLMC-FVM

The MLMC FVM consists in estimates of $\mathbb{E}[u(\cdot, t)]$ obtained by replacing $u(\cdot, t)$ by a FV discretization, on a sequence of discretizations $\{\mathcal{T}_\ell\}_{\ell \in \mathbb{N}_0}$ which we assume to be nested. We denote in the present section the FV approximation $v_{\mathcal{T}}$ on triangulation $\mathcal{T} \in \mathfrak{M}$ by $v_\ell(\cdot, t)$. On $\mathcal{T}_\ell \in \mathfrak{M}$, the CFL condition (58) takes the form

$$\Delta t_\ell \leq \lambda \Delta x_\ell, \quad \ell = 0, 1, 2, \dots, \quad (89)$$

We assume the CFL constant $\lambda > 0$ to be independent of ℓ and of the input realization ω ; this will allow for deterministic error vs. work bounds; we refer to [49] for a discussion of error vs. work of MLMC for nonuniform (log-gaussian) random inputs.

As the FV scheme is CFL stable, we may generate a sequence $\{v_\ell(\cdot, t)\}_{\ell=0}^\infty$ of stable FV approximations on triangulation \mathcal{T}_ℓ for time steps of sizes Δt_ℓ which satisfy the CFL condition (89) with respect to mesh $\mathcal{T}_\ell \in \mathfrak{M}$. We set in what follows $v_{-1}(\cdot, t) := 0$. Then, given a target (finest) level $L \in \mathbb{N}$ of spatial resolution, we may use the linearity of the expectation operator to write, as is customary in MLMC analysis (see, e.g., [27])

$$\mathbb{E}[v_L(\cdot, t)] = \sum_{\ell=0}^L \mathbb{E}[v_\ell(\cdot, t) - v_{\ell-1}(\cdot, t)]. \quad (90)$$

We next estimate each term in the sum (90) by a Monte-Carlo method with a level-dependent number of samples, M_ℓ , to obtain the MLMC-FVM estimator,

$$E^L[u(\cdot, t)] = \sum_{\ell=0}^L E_{M_\ell}[v_\ell(\cdot, t) - v_{\ell-1}(\cdot, t)]. \quad (91)$$

Here, $E_M[v_{\mathcal{T}}(\cdot, t)]$ is the standard MC estimator defined in (79), and $v_\ell(\cdot, t)$ denotes the FV solution on \mathcal{T}_ℓ , computed under the CFL assumption (89), with $\Delta t_\ell \leq \lambda \Delta x_\ell$ where $\Delta x_\ell := \Delta x(\mathcal{T}_\ell)$ denotes the meshwidth at mesh level ℓ (see (55)) and where the CFL constant $\lambda > 0$ is independent of ℓ . We emphasize that the form of the estimator (90) implies that the same draw of the random initial data should be approximated on two successive meshes in the hierarchy.

4.4.3 Convergence Analysis

The MLMC-FVM mean field error

$$\left\| \mathbb{E}[u(\cdot, t)] - E^L[u(\cdot, t)] \right\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \quad (92)$$

for $0 < t < \infty$ and $L \in \mathbb{N}$ was analyzed in [47] for the SCL (13) with random initial data and deterministic flux. Analogous results for the more general SCL with random fluxes was shown in [46]. The choice of the sample sizes $\{M_\ell\}_{\ell=0}^\infty$ is such that, for every $L \in \mathbb{N}$, the MLMC error (92) is of order $(\Delta x_L)^s$, where s is the order of convergence in the Kuznetsov type error bound (68). The design of MLMC-FVM is based on a judicious choice of MC sample numbers $\{M_\ell\}_{\ell=0}^\infty$ at the discretization levels ℓ . To derive it, we observe that for each L , the error bound (92) holds with work bounded by

$$\text{Work}_L = \sum_{\ell=0}^L M_\ell \mathcal{O}(\Delta x_\ell^{-d-1}) = \mathcal{O}\left(\sum_{\ell=0}^L M_\ell \Delta x_\ell^{-d-1}\right). \quad (93)$$

The MLMC convergence analysis in [47] used incorrectly a MC convergence estimate for the space $L^1(\mathbb{R}^d)$ which is a Banach space of type 1. This error was rectified in a recent paper [46], where the following bound on the variance of the FV details was shown:

$$\|(v_\ell - v_{\ell-1})(\cdot, t)\|_{L^2(\Omega; L^1(\mathbb{R}^d))}^2 \leq C(D, T) \Delta x_\ell^s \|u_0\|_{L^2(\Omega; W^{s,2}(\mathbb{R}^d))}^2. \quad (94)$$

Theorem 5. [46, Thm. 4.7] *Suppose that Assumption 1, items 1. - 5. hold, and that, moreover, (28) - (32) and (87) - (89) are valid. Then, for any sequence $\{M_\ell\}_{\ell=0}^\infty$ of MC sample numbers at mesh level ℓ , we have for the MLMC-FVM estimate $E^L[u(\cdot, t)]$ in (91) the error bound*

$$\begin{aligned} & \left\| \mathbb{E}[u(\cdot, t)] - E^L[u(\cdot, t)] \right\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \\ & \leq C(D, T, \|u_0\|_{L^\infty(\Omega; Z)}) \left[\Delta x_L^s + \sum_{\ell=1}^L M_\ell^{-\frac{1}{2}} \Delta x_\ell^{\frac{s}{2}} + M_0^{-\frac{1}{2}} \right] \end{aligned} \quad (95)$$

where C is a constant that depends on the final time $T < \infty$, the initial data and on the bounded domain D which contains, according to (52), the support of $u(t)$ with probability one, but is independent of L .

Proof. We calculate for any $t \in [0, T]$

$$\begin{aligned} \left\| \mathbb{E}[u(\cdot, t)] - E^L[u(\cdot, t)] \right\|_{L^2(\Omega; L^1(\mathbb{R}^d))} & \leq \underbrace{\left\| \mathbb{E}[u(\cdot, t)] - \mathbb{E}[v_L(\cdot, t)] \right\|_{L^2(\Omega; L^1(\mathbb{R}^d))}}_I \\ & \quad + \underbrace{\left\| \mathbb{E}[v_L(\cdot, t)] - E^L[u(\cdot, t)] \right\|_{L^2(\Omega; L^1(\mathbb{R}^d))}}_{II} \end{aligned}$$

In complete analogy with the estimation of term II in proof of Theorem 4, the term I in the above estimate can be readily estimated in terms of the spatio-temporal discretization error of the FV scheme at the finest mesh resolution with diameter Δx_L as follows,

$$I \leq C(T, \|u_0\|_{L^\infty(\Omega;Z)}) \Delta x_L^s.$$

To estimate term II in the above estimate, we use the discrete finite dependence domain result, Proposition 3, and the bounded support assumption (32) on the random initial data and proceed as follows,

$$\begin{aligned} II &= \left\| \mathbb{E}[v_L(\cdot, t)] - E^L[u(\cdot, t)] \right\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\ &= \left\| \sum_{l=0}^L \left[\mathbb{E}[v_l(\cdot, t) - v_{l-1}(\cdot, t)] - E_{M_\ell}[v_\ell(\cdot, t) - v_{\ell-1}(\cdot, t)] \right] \right\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\ &\leq \sum_{l=0}^L \left\| \mathbb{E}[v_l(\cdot, t) - v_{l-1}(\cdot, t)] - E_{M_\ell}[v_\ell(\cdot, t) - v_{\ell-1}(\cdot, t)] \right\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\ &\leq C(D, T) \sum_{l=0}^L \left\| \mathbb{E}[v_l(\cdot, t) - v_{l-1}(\cdot, t)] - E_{M_\ell}[v_\ell(\cdot, t) - v_{\ell-1}(\cdot, t)] \right\|_{L^2(\Omega;L^2(\mathbb{R}^d))} \\ &\leq C(D, T) \left\{ \frac{\|v_0(\cdot, t)\|_{L^2(\Omega;L^2(\mathbb{R}^d))}}{M_0^{\frac{1}{2}}} + \sum_{l=1}^L \left(\frac{\|(v_\ell - v_{\ell-1})(\cdot, t)\|_{L^2(\Omega;L^2(\mathbb{R}^d))}}{M_l^{\frac{1}{2}}} \right) \right\} \end{aligned}$$

In the final estimate, we used the compact support assumption (50) and in the final step, we used the standard Hilbert space MC estimate (12) for the *detail* $v_\ell - v_{\ell-1}$. Accordingly, we need to bound the variance of the details $v_\ell - v_{\ell-1}$ in the (Hilbert) space $L^2(\mathbb{R}^d)$ according to

$$\|(v_\ell - v_{\ell-1})(\cdot, t)\|_{L^2(\Omega;L^2(\mathbb{R}^d))} \leq C(D, T) \|u_0\|_{L^2(\Omega;Z)} \Delta x_\ell^{\frac{s}{2}}. \quad (96)$$

Note that the use of the (Hilbertian) $L^2(\mathbb{R}^d)$ norm in the error bound entails the convergence rate $s/2$ of the FV detail $v_\ell - v_{\ell-1}$, where $0 < s \leq 1$ denotes the $L^1(\mathbb{R}^d)$ convergence rate of the (deterministic) FV approximation. Substituting the above in the estimate for term I and using the stability of the numerical solution at the coarsest level of discretization Δx_0 , we arrive at (95). \square

The error bound (95) is then used to select MC sample numbers M_ℓ at discretization level ℓ to achieve a prescribed tolerance $\varepsilon \sim \Delta x_L^s$, with minimal computational work.

A (standard in MLMC by now) Lagrange multiplier argument (see, e.g., Giles [27] and references therein) allows to solve the corresponding constrained minimization problems results in sample number choices obtained, for example, in [46] for $0 < s < d + 1$,

$$M_\ell \sim \frac{\Delta x_\ell^{\frac{(s+d+1)}{2}}}{\Delta x_L^{2s}} \sum_{k=0}^L \Delta x_k^{\frac{(s-(d+1))}{2}} \sim \frac{\Delta x_\ell^{\frac{(s+d+1)}{2}}}{\Delta x_L^{\frac{(3s+d+1)}{2}}} \quad (97)$$

with \sim denoting equivalence uniform with respect to L and ℓ .

As in [47], we use the sample numbers M_ℓ in (97) to obtain the following error estimate in terms of work

$$\|\mathbb{E}[u(\cdot, t)] - E^L[u(\cdot, t)]\|_{L^2(\Omega; L^1(\mathbb{R}^d))} \leq C (\text{Work}(\{M_\ell\}_{\ell=0}^L; \mathcal{F}_L))^{-s/(d+1+s)}. \quad (98)$$

The complexity estimate (98) shows that the MLMC FVM can be more efficient than the MC FVM (84), in terms of computational work that needs to be performed for obtaining the same error. However, to achieve a comparable error in $L^2(\Omega; L^1(\mathbb{R}^d))$, the MLMC method is more expensive than a single deterministic solve.

Remark 4. The above discussion on random entropy solutions and (ML)MC methods considered the simplest case of a scalar conservation law with random initial data. These notions and methods were extended to the case where the flux function in a scalar conservation law is random. In a recent paper [46], where the appropriate notion of random entropy solutions were defined and shown to exist, provided that the uncertainty in the flux satisfied certain assumptions, which ensure the random flux to be Bochner measurable and \mathbb{P} -a.s. separably valued in the space of Lipschitz continuous flux functions. Both the MC-FV and MLMC-FV methods were analyzed in this context and the MLMC-FV method was shown to satisfy the same error vs computational work i.e, (98) as in the case of deterministic fluxes and random data. Consequently, the MLMC method is significantly more efficient than the corresponding MC-FV method. We refer to [46] for details.

5 Statistical Moments

The error bounds for the MLMC-FVM obtained up to this point addressed the numerical estimation of the “ensemble average”, or mean-field. Here, we briefly comment on efficient numerical approximation of 2- and k -point correlation functions of random solutions. When the physical problem is posed in d spatial variables, *spatial k -point correlation functions* of random or of statistical solutions can be represented (under the provision of sufficient regularity) as deterministic functions of kd variables.

The “natural”, FV approximation of k -point spatial correlations of random entropy solutions as well as of correlation margins of statistical EMV solutions introduced in [19, 16], will involve k -fold (algebraic) tensorisation of (finite dimensional, by the bounded support assumption (32) on the random initial data, and by the corresponding bounded support property (71) of FV approximations implied by the uniform hyperbolicity of the SCL). This can increase computational complexity due to the low convergence rate $1/2$ of MC sampling, and to the so-called “curse

of dimensionality”, which entails complexity $\mathcal{O}(\Delta x^{-kd-1})$ in k -point correlation estimation. This can be prohibitive, in particular in space dimension $d = 3$, even for two-point correlations where $k = 2$. Two algorithmic strategies are next presented which allow, to some extent and under appropriate regularity conditions, to reduce the computational complexity: first, the MC statistical estimation of k -th central statistical moments in the physical dimension D as analyzed recently in [3], and second, k -fold sparse tensor products of FV solutions in the physical domain $D \subset \mathbb{R}^d$ as proposed in [47]. They render k -point correlations in $D^k \subset \mathbb{R}^{kd}$ numerically accessible in $\mathcal{O}(\Delta x^{-d-1} |\log \Delta x|^{k-1})$ operations.

5.1 Estimation of k -th order central statistical moments

Given a random entropy solution u , a “natural” MC estimator based on M iid samples for the k -th central moment of $u(\cdot; x, t)$ reads

$$S_M^k[u] := \frac{1}{M} \sum_{i=1}^M (u^i - E_M[u])^k, \quad (99)$$

assuming availability of exact solution samples u^i of the random entropy solution. The estimator (99) is to be interpreted pointwise w.r. to $x \in D$ and $t \in [0, T]$. It is, however, well-known to be a statistically biased estimator. Unbiased estimators are known. For example,

$$\tilde{S}_M^2[u] := \frac{M}{M-1} S_M^2[u], \quad \tilde{S}_M^3[u] := \frac{M^2}{(M-1)(M-2)} S_M^3[u] \quad (100)$$

are unbiased estimators of $\mathcal{M}^2(u)$ and $\mathcal{M}^3(u)$. For $k \geq 4$, unbiased estimators $\tilde{S}_M^k[u]$ can be obtained as polynomial expressions of $S_M^r[u]$ for $r = 1, \dots, K$ which are, however, not unique. We refer to [3, Lemma 3] for details.

A MLMC estimator of $\mathcal{M}^k(u)(x, \dots, x; t)$ is introduced and analyzed in [3, Theorem 1].

As the corresponding algorithms only access the FV solver through iid samples of the random initial data and random flux, respectively, they are nonintrusive and embarrassingly parallel. Operating only in the physical domain $D \subseteq \mathbb{R}^d$, they do not require additional data structures for tensorization of FV solutions. Central statistical moments $\mathcal{M}^k(u)(x, \dots, x; t)$ can, however, also be numerically approximated as “diagonals” of statistical k -point correlations $\mathcal{M}^k(u)$. We discuss this next.

5.2 Estimation of k -point spatial correlations

The work to form a single tensor product over a bounded computational domain $D \subset \mathbb{R}^d$ (such as, e.g., the support domain of the solution at time t in (26)) grows, ignoring timestepping, as $O(\Delta x^{-kd})$ which may entail a computational effort that is, for moment orders $k \geq 2$, prohibitive. To reduce the complexity of k -point correlation function estimation, a so-called “sparse grid” or “sparse tensor product” approach was proposed in [52, 2, 47]. We refer to [60, 31, 56] and the references there for general presentations of sparse tensor product spaces.

5.2.1 k -point correlation estimation by sparse tensorization of FV Solutions

As is standard in multi-resolution analysis, the cell-average projections $P_\ell : L^1(\mathbb{R}^d) \rightarrow S_\ell$, defined in (88), (104), allow us to introduce spaces of *increments* or *details* in the FV mesh hierarchy $\mathfrak{M} = \{\mathcal{T}_\ell\}_{\ell=0}^\infty$:

$$W_\ell := (P_\ell - P_{\ell-1})S_\ell, \quad \ell \geq 0 \quad (101)$$

where $P_{-1} := 0$ so that $W_0 = S_0$. Then, for any $L \in \mathbb{N}_0$, we have the multilevel decomposition

$$S_L = W_0 \oplus \dots \oplus W_L = \bigoplus_{\ell=0}^L W_\ell. \quad (102)$$

The k -point correlations $(v_L(\cdot, t))^{(k)}$ of FV solutions at time $t > 0$ belong to the (algebraic) tensor product space

$$(S_L)^{(k)} := \underbrace{S_L \otimes \dots \otimes S_L}_{k \text{ times}} = \sum_{|\ell|_\infty \leq L} S_{\ell_1} \otimes \dots \otimes S_{\ell_k} = \bigoplus_{|\ell|_\infty \leq L} \bigotimes_{j=1}^k W_{\ell_j}. \quad (103)$$

In the numerical realization of 2- and k point correlation functions and correlation margins of measure valued solutions, the computational realization of approximations in the tensor product space $(S_L)^{(k)}$ is necessary. To formulate it, we introduce the k -fold algebraic tensor products of the FV cell-average projections by

$$P_L^{(k)} v := \underbrace{P_L \otimes \dots \otimes P_L}_{k \text{ times}} : L^1(\mathbb{R}^{kd}) \rightarrow (S_L)^{(k)}. \quad (104)$$

In the case that $N_L := \dim S_L < \infty$ (as e.g. when the spaces S_ℓ are only defined on a bounded domain $D \subset \mathbb{R}^d$ such as the “support box” (26) in Prop. 2) then $\dim((S_L)^{(k)}) = N_L^k$ which is prohibitive. *Sparse Tensor approximations of k -point correlation functions*, $(v(\cdot, t))^{(k)}$ are “compressed” FV approximations which involve FV spaces of piecewise constant functions on coarser meshes. They are defined in terms of the *increment space* W_ℓ in (101) by

$$\widehat{(S_L)}^{(k)} := \bigoplus_{|\underline{\ell}|_1 \leq L} \bigotimes_{j=1}^k W_{\ell_j} \quad (105)$$

where $|\underline{\ell}|_1 := \ell_1 + \dots + \ell_k$ and where algebraic tensor products are implied. Note that a realization of the sparse tensor product space $\widehat{(S_L)}^{(k)}$ according to its definition (105) requires construction of explicit bases for W_{ℓ} in (101); on unstructured, simplicial triangulations \mathcal{T} as in our Assumption 2 such bases can be numerically constructed by agglomeration, see, e.g. [1]. A one-scale FV solution on the finest mesh \mathcal{T}_L can be converted to a ML representation in $O(N_L)$ operations by the so-called *pyramid scheme* (see, e.g., [5, Pp. 225-294]). If the mesh family \mathfrak{M} used in the pathwise FV approximation (see Assumption 2) is generated by recursive dyadic refinements of the initial triangulation \mathcal{T}_0 , when $N_L = \dim S_L < \infty$ (as is the case e.g. on bounded domains $D \subset \mathbb{R}^d$) it holds (see, e.g. [60, 56])

$$\dim \widehat{(S_L)}^{(k)} = O(N_L (\log_2 N_L)^{k-1}). \quad (106)$$

Having at hand the sparse tensor product space $\widehat{(S_L)}^{(k)}$ in (105), we also define the *sparse tensor projection*

$$\widehat{(P_L)}^{(k)} := \bigoplus_{|\underline{\ell}|_1 \leq L} \bigotimes_{j=1}^k (P_{\ell_j} - P_{\ell_j-1}) : L^1(\mathbb{R}^{kd}) \rightarrow \widehat{(S_L)}^{(k)}. \quad (107)$$

We refer to [31, 60] and the references there, from where we briefly recapitulate approximation properties of sparse tensor product projections: for any function $U(x_1, \dots, x_k)$ which belongs to $(W^{s,1}(\mathbb{R}^d))^{(k)}$ being the space of functions of k variables $x_1, \dots, x_d \in \mathbb{R}^d$ which are, with respect to each variable, in $W^{s,1}(\mathbb{R}^d)$, it holds

$$\|U - \widehat{(P_L)}^{(k)} U\|_{L^1(\mathbb{R}^{kd})} \leq C(\Delta x_L)^s |\log \Delta x_L|^{k-1} \|U\|_{(W^{s,1}(\mathbb{R}^d))^{(k)}} \quad (108)$$

where $C > 0$ is independent of Δx_L (it depends only on k, d and the shape regularity of the family \mathfrak{M} of triangulations, but is independent of Δx_L).

5.2.2 Sparse MLMC-FVM Estimator

The MLMC sparse FV estimator is based on a sparse tensor product FV approximation for each MC sample. To define it, we recall that $E_M[\cdot]$ denotes the MC estimate based on M samples): for a given sequence $\{M_\ell\}_{\ell=0}^L$ of MC sample numbers at level $\ell = 0, \dots, L$, the sparse tensor MLMC estimate of $\mathcal{M}^k[u(\cdot, t)]$ is, for $0 < t < \infty$, defined by

$$\widehat{E}^{L,(k)}[u(\cdot, t)] := \sum_{\ell=0}^L E_{M_\ell}[\widehat{P}_\ell^{(k)}(v_\ell(\cdot, t))^{(k)} - \widehat{P}_{\ell-1}^{(k)}(v_{\ell-1}(\cdot, t))^{(k)}]. \quad (109)$$

We observe that (109) is identical to (103) except for the sparse formation of the k -point correlation functions of the FV solutions corresponding to the initial data samples \hat{u}_0^i . In bounded domains, this reduces the work for the formation of the k -point correlation function from N_L^k to $O(N_L(\log_2 N_L)^{k-1})$ per sample at mesh level L . As is well known (see, e.g. [31, 60]) use of sparse rather than full tensor products essentially preserves (ie., up to logarithmic w.r. to Δx terms) the order s of FV convergence of sparse tensor product k point correlation function approximations.

5.2.3 Combination Formula

Sparse tensor products are particularly easy to realize when the FV scheme already produces FV solutions to SCLs in MRA format. Such schemes are nonstandard, but available even on unstructured meshes as we admit in Assumption 2, in a development [50, 25] originating in the seminal work of A. Harten [4, 1]. Often, however, only single-level (one-scale) numerical FV approximations on triangulations \mathcal{T}_ℓ , $\ell = 0, \dots, L$ are available. In order to realize the MLMC-FV estimator (109), for each realization the approximations must be converted to a MR representation. This can be achieved also on unstructured meshes in linear complexity by the so-called *pyramid scheme* (see, e.g., [55] for a definition and an algorithm).

An alternative approach which obviates MR based numerical methods is the so-called *combination formula*, as proposed for this purpose (in a different context) in [32, Lemma 12, Thm. 13]. For the projector $(\widehat{P}_L)^{(k)}$ in (107), the *combination identity*

$$(\widehat{P}_L)^{(k)} = \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} \sum_{|\underline{\ell}|_1=k-i} P_{\underline{\ell}}^{(k)}, \quad \text{where } P_{\underline{\ell}}^{(k)} := \bigotimes_{j=1}^k P_{\ell_j} \quad (110)$$

holds. The combination identity (110) implies that sparse tensor MC-FV approximations of k -point correlation functions can be numerically built from (pointwise) products of standard, one-scale FV approximations of the SCL (13) - (15) with iid samples of the random initial data u_0 on mesh levels $\underline{\ell} := (\ell_1, \dots, \ell_k)$. When inserted into (109), the combination identity (110) provides an explicit realization of the MLMC-FVM estimator of k -point correlations of random entropy solutions, based exclusively on (parallel) standard FV solves on all mesh levels.

5.2.4 Error Bounds and Complexity Analysis

We can now generalize Theorems 4 and 5 to sparse tensor MLMC-FV estimates for the k -th moments of random entropy solutions; it also applies to estimates of correlation measures of entropy statistical solutions as introduced in [17].

Theorem 6. *Assume that the random initial data u_0 satisfies Assumption 1, items 1.-4. and that the FV scheme satisfies Assumption 2, items 1.-3.. In particular, assume*

that the deterministic FV scheme converges at rate $s > 0$ according to (68) and that for u_0 , the bounded support Assumption 2, item 4., (32) holds.

Assume further that the FV scheme satisfies the CFL condition (58) and that the random initial data u_0 is Bochner-integrable of order $2k$ in the spaces $Z = (L^1 \cap L^\infty \cap BV)(\mathbb{R}^d)$ and in $W^{s,1}(\mathbb{R}^d)$, i.e.

$$u_0 \in L^{2k}(\Omega; Z) \cap L^{2k}(\Omega; W^{s,1}(\mathbb{R}^d)). \quad (111)$$

Then, MLMC-FVM estimator $\widehat{E^{L,(k)}}[u(\cdot, t)]$ defined in (109) satisfies, for every sequence $\{M_\ell\}_{\ell=0}^L$ of MC samples, the error bound

$$\begin{aligned} & \|\mathcal{M}^k u(\cdot, t) - \widehat{E^{L,(k)}}[u(\cdot, t; \omega)]\|_{L^2(\Omega; L^1(\mathbb{R}^{kd}))}^2 \\ & \lesssim (1 \vee t) \Delta x_L^{2s} |\log \Delta x_L|^{2(k-1)} \left\{ \|\text{TV}(u_0(\cdot, \omega))\|_{L^k(\Omega; d\mathbb{P})}^{2k} + \|u_0(\cdot; \omega)\|_{L^\infty(\Omega; W^{s,1}(\mathbb{R}^d))}^{2k} \right\} \\ & + \left\{ \sum_{\ell=0}^L \frac{\Delta x_\ell^s |\log \Delta x_\ell|^{k-1}}{M_\ell} \right\} \left\{ \|u_0(\cdot; \omega)\|_{L^{2k}(\Omega; W^{s,1}(\mathbb{R}^d))}^{2k} + t^2 \|\text{TV}(u_0(\cdot; \omega))\|_{L^{2k}(\Omega; d\mathbb{P})}^{2k} \right\}. \end{aligned}$$

Here, the constant implied in \lesssim depends on the order k of the moment to be estimated, on the physical space dimension d , and on the support size constant $R > 0$ in the bounded support assumption (32).

Then, the total work to compute the MLMC estimates $\widehat{E^{L,(k)}}[u(\cdot; t)]$ is bounded by (with $O(\cdot)$ depending on the size of D)

$$\widehat{\text{Work}}_L^{MLMC} = O\left(\sum_{\ell=0}^L M_\ell \Delta x_\ell^{-(d+1)} |\log \Delta x_\ell|^{k-1}\right). \quad (112)$$

Based on Theorem 6, we infer that the choice (97) of numbers M_ℓ of MC samples at level ℓ should also be used in the MLMC estimation of k -point correlation functions for $k > 1$, provided the order s of the underlying deterministic FVM scheme (59) - (61) satisfies

$$0 \leq s < d + 1. \quad (113)$$

In particular, in a bounded domain $D \subset \mathbb{R}^d$ containing the bounded support in (52),

$$\|\mathcal{M}^k u(\cdot, t) - \widehat{E^{L,(k)}}[u(\cdot, t; \omega)]\|_{L^2(\Omega; L^1(D^k))} \leq C(\widehat{\text{Work}}_L^{MLMC})^{-s'/(d+1+s)} \quad (114)$$

for any $0 < s' < s$ (with constant growing as $0 < s' \rightarrow s \leq 1$). The computational domain D can, in particular, contain the bounded support domains of the exact and discrete solutions at time $t > 0$, as indicated in (71), (52).

6 Monte Carlo and Multi-level Monte Carlo methods for systems of conservation laws

6.1 General considerations

We consider the general system of conservation laws (1), with random initial data in (115b) as well as possibly random coefficients and random flux functions in (115a). A notion of random entropy solutions can be defined for this general case, analogous to definition 2 for scalar conservation laws. We refer to [48] for details. However, there are no well-posedness results for random entropy solutions for systems as even the underlying deterministic problem may not be well-posed, particularly in several space dimensions [14]. One approach to developing numerical approximations in this case is to *assume existence of random entropy solutions* and to design efficient methods for numerical approximations of their solutions:

It is fairly straightforward to extend the MCFV scheme, given in section 4.3.1, to general, nonlinear hyperbolic systems of conservation laws with random inputs. A convergence rate estimate, analogous to (80) can be proved, *provided* that one postulates a convergence rate, analogous to (68), for the underlying spatio-temporal FV discretization, see [48]. Similarly, the MLMC method, as described in section 4.4 can also be readily extended to this general case and a convergence rate estimate, similar to (95), once the underlying spatio-temporal discretization converges like in (68) or at least a estimate of the type (96) holds. Consequently, a complexity estimate as (98) can be shown under these assumptions, demonstrating that the MLMC-FV method is more efficient than the MC-FV method.

6.2 Numerical experiments

We present a few numerical experiments involving systems of conservation laws to illustrate the robustness of the MLMC method and its comparison with the Monte Carlo method.

6.2.1 Uncertain Orszag-Tang vortex

This test is taken from [48], section 6.2. The system of conservation laws under consideration are the ideal magnetohydrodynamics (MHD) equations of plasma physics. We consider the ideal MHD equations on the two-dimensional domain $[0, 1]^2$ with periodic boundary conditions. The random initial data is an uncertain version of the well-known Orszag-Tang benchmark test problem. We consider a random initial data with 8 sources of uncertainty, namely in the amplitudes of the initial density and pressure, phases of the initial sinusoidal velocity fields and the phases and amplitude of the initial solenoidal magnetic fields. The mean and the variance of the density,

computed at time $T = 1.0$ with an MLMC-FV method, with 8 levels of resolution, a finest mesh of 4096^2 and with 4 samples at the finest level, with the underlying FV method using a HLLC Riemann solver, a second-order WENO reconstruction and upwind treatment of the Godunov-Powell source term [23], are shown in figure 1. In this case, one computes a reference solution with the above configuration and calculates the $L^2(\Omega, L^1(D))$ error for both the mean and variance, with MC and MLMC methods (of both first and second order spatio-temporal discretizations). The errors for the mean and variance are plotted in figures 2 and 3, respectively. They show that in this example, the MLMC FV method is at least 50 – 60 times faster than the single-level MC FV method for the mean and 10 – 20 times faster for the variance, to achieve a prescribed error tolerance in the engineering range of accuracy. Thus, this justifies the complexity estimates, described here, at least for this example.

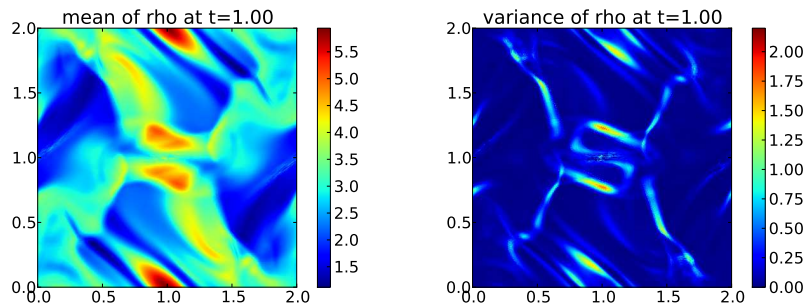


Fig. 1 Uncertain Orszag-Tang vortex solution at $t = 1.0$ using MLMC-FVM (8 sources of uncertainty). Left: Convergence of the sample mean of random density. Reproduced from [59]

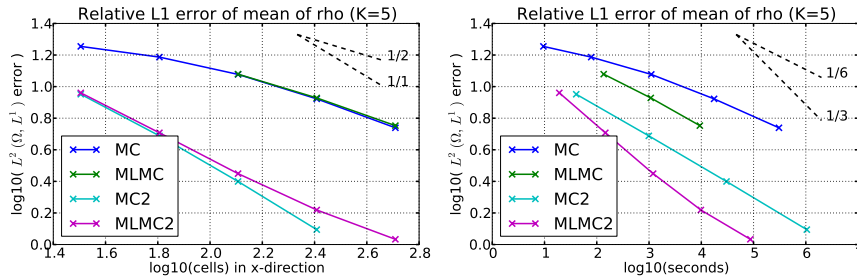


Fig. 2 Convergence of sample mean in the uncertain Orszag-Tang vortex simulation (8 sources of uncertainty). Left: Error vs. Mesh resolution. Right: Error vs. Runtime. Reproduced from [59].

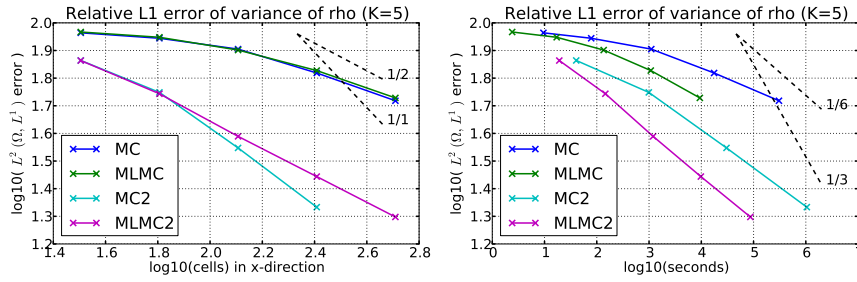


Fig. 3 Convergence of sample variance in the uncertain Orszag-Tang vortex simulation (8 sources of uncertainty). Left: Error vs. Mesh resolution. Right: Error vs. Runtime. Reproduced from [59].

6.2.2 A random Kelvin-Helmholtz problem

This numerical example is taken from a recent paper [16] and the MLMC computations are presented in [45]. We consider the compressible Euler equations in the two-dimensional domain $[0, 1]^2$ with periodic boundary conditions. The uncertainty arises due to the initial data being a (very small) random perturbation of the classic Kelvin-Helmholtz problem, see [16], with 20 sources of uncertainty in the initial data. The mean and variance of the density, computed with a Monte Carlo method, on a Cartesian 1024^2 grid and with 400 MC samples are shown in Figure 4 (Top Row). The underlying FV scheme is the third-order entropy stable TeCNO scheme of [18]. Surprisingly for this test case, the variance of the solution is at least three orders of magnitude higher than the variance of the initial data. This amplification of variance is due to the generation of structures at smaller and smaller scales, when the shear flow interacts with the contact discontinuity. In this particular case, the MLMC method provides *no gain* in computational efficiency over the standard Monte Carlo method. This is depicted in Figure 4 (Bottom row, Left), where the L^1 difference in the mean of the density, computed with the MLMC and MC methods, at the same grid resolution for the finest grid and the same number of samples at the finest grid resolution of MLMC, with respect to an MC reference solution computed with 1024 samples, is compared. The results show that error due to MLMC is comparable to the error due to the MC calculation, provided that the number of samples at the finest grid level of the MLMC calculation is the same as the number of MC samples. Thus, in this case, the coarse levels of the MLMC method do not increase the accuracy of the computation and are redundant. Given this observation, it is clear that an error estimate of the form (95) cannot hold for this particular example. In fact, even an error estimate of the form (68) for the underlying spatio-temporal discretization, does not hold for this example. We see this from Figure 4 (Bottom row, Right) where the difference in L^1 between two successive mesh resolutions for a single sample is shown. This figure shows that the error remains constant with respect to resolution and the underlying FV scheme does not converge for this particular test case. It was also shown in [45] that even weaker convergence bounds, such as the variance of

the difference between successive resolutions (96), does not hold for this particular problem, as structures at even smaller scales are generated upon mesh refinement.

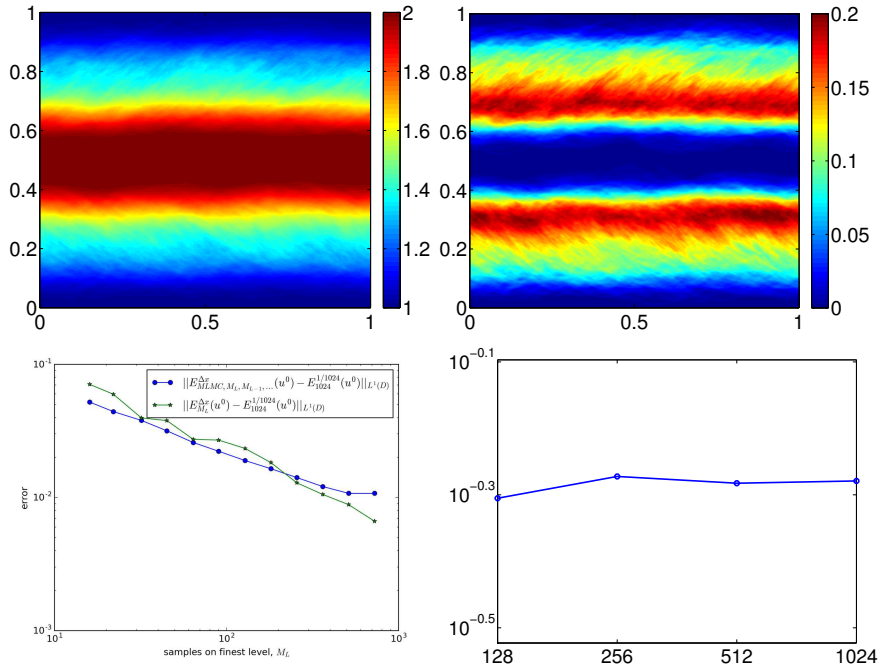


Fig. 4 UQ for the random Kelvin-Helmholtz problem. Top Row: Density at time $T = 2$ computed with MC method and TeCNO3 scheme with the algorithm proposed in [16] Left: Convergence of the sample mean, Right: Convergence of the sample variance. Bottom Row Left: Comparison between MC and MLMC method with respect to error in mean of the density with respect to a reference solution (Reproduced from [45]). Bottom Right: Lack of convergence for a single sample for the density (reproduced from [45]).

7 Measure-valued and Statistical solutions

It is interesting to note that the Monte Carlo and Multi-level Monte Carlo methods converge for the previous numerical experiment, as shown in Figure 4 (bottom left), even though the underlying spatio-temporal discretization may not converge in L^1 , as shown in Figure 4 (bottom right). What exactly do the (ML)MC-FV computations converge to in this case ?

7.1 Measure-valued solutions

This question was partially answered in recent papers [16, 19] and references therein. There, the authors proved that a Monte Carlo based algorithm, together with an entropy stable scheme such as the TeCNO scheme of [18], converges to an *entropy measure valued solution* of the underlying system of conservation laws (1). Measure valued solutions were first proposed by [53] and are in fact *Young measures* i.e., space-time parametrized probability measures on the phase space \mathbb{R}^N of the system (1). Let $D \subset \mathbb{R}^d$ be the domain and $D_t := D \times \mathbb{R}_+$, we define *Young measure* from D_t to \mathbb{R}^N as a map which associates to each point $(x, t) \in D_t$ a probability measure on \mathbb{R}^N . More precisely, a Young measure is a weak* measurable map $\nu : D_t \rightarrow \mathcal{P}(\mathbb{R}^N)$, meaning that the mapping

$$(x, t) \mapsto \langle \nu_{x,t}, g \rangle = \int_{\mathbb{R}^N} g(\xi) d\nu_{x,t}(\xi) \text{ is Borel measurable for every } g \in C_0(\mathbb{R}^N).$$

The set of all Young measures from D_t into \mathbb{R}^N is denoted by $\mathbf{Y}(D_t, \mathbb{R}^N)$. Given this notation for Young measures, one can rewrite the following system of N -conservation laws,

$$\partial_t u + \nabla_x \cdot f(u) = 0, \quad (x, t) \in D_t, \quad (115a)$$

$$u(x, 0) = \bar{u}(x), \quad x \in D. \quad (115b)$$

in terms of the following *measure-valued Cauchy problem*,

$$\begin{aligned} \partial_t \langle \nu_{x,t}, \xi \rangle + \nabla_x \cdot \langle \nu_{x,t}, f(\xi) \rangle &= 0, \quad (x, t) \in D_t, \\ \langle \nu_{x,0}, \xi \rangle &= \langle \sigma_x, \xi \rangle, \quad x \in D, \end{aligned} \quad (116)$$

with possibly Young measure-valued initial data σ_x . The above system (116) has to hold in the sense of distributions. Entropy (admissibility) conditions can be imposed by interpreting an associated entropy inequality in the Young measure sense [53]

Global existence of measure-valued solutions was shown recently in [16, 19] and references therein, by proving convergence of the following Monte Carlo based ensemble averaging algorithm,

Algorithm 1

Let the initial data for an underlying time-dependent PDE (116) be given as a Young measure $\sigma \in \mathbf{Y}(D, \mathbb{R}^N)$ i.e., a Young measure $D \mapsto \mathcal{P}(\mathbb{R}^N)$.

- **Step 1:** Let $u_0 : \Omega \mapsto L^p(\mathbb{R}^d)$ be a random field on a probability space $(\Omega, \Sigma, \mathbb{P})$ with law σ , i.e., $\sigma(E) = \mathbb{P}(u_0(\omega) \in E)$.
- **Step 2:** Evolve the initial random field by applying a suitable numerical scheme, with solution map \mathcal{S}_t^Δ , to the initial data $u_0(\omega)$ for every $\omega \in \Omega$, obtaining an approximate random field $u^\Delta(\omega; \cdot, t) := \mathcal{S}_t^\Delta u_0(\omega; \cdot)$.
- **Step 3:** Define the *approximate measure-valued solution* ν^Δ as the law of u^Δ with respect to \mathbb{P} , i.e. for all Borel sets $E \subset \mathbb{R}^N$,

$$v_{x,t}^\Delta(E) = \mathbb{P}(u(\omega; x, t) \in E).$$

It was shown in [16, Appendix A.3.1] that v^Δ are indeed Young measures. The existence of a random field u_0 with a given law σ , as required in Step 1, is guaranteed by proposition A.3 of [16].

The numerical method in Step 2 of algorithm 1, can be appropriate structure preserving Finite Volume Methods, such as the arbitrary high-order entropy stable TeCNO schemes of [18]. The last ingredient in our numerical approximation of measure-valued solutions is to find, and approximate, the random field $u_0(\omega; x)$ which appears in Algorithm 1, resulting in the following algorithm.

Algorithm 2.

Let $\Delta = (\Delta x_1, \dots, \Delta x_d)$ denote the grid size parameter and let $M \in \mathbb{N}$. Let further $\sigma^\Delta \in \mathbf{Y}(\mathbb{R}^d, \mathbb{R}^N)$ denote the initial Young measure.

- **Step 1:** For some probability space $(\Omega, \Sigma, \mathbb{P})$, draw M i.i.d. random fields $u_0^{\Delta,1}, \dots, u_0^{\Delta,M} : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^N$, with [the same] law σ^Δ .
- **Step 2:** For each k and for a fixed $\omega \in \Omega$, use a suitable numerical scheme to numerically approximate the conservation law (115a) with initial data $u_0^{\Delta,k}(\omega)$. Denote $u^{\Delta,k}(\omega; \cdot, t) = \mathcal{S}_t^\Delta u_0^{\Delta,k}(\omega; \cdot)$ as the computed solutions.
- **Step 3:** Define the approximate measure-valued solution

$$v_{x,t}^{\Delta,M} := \frac{1}{M} \sum_{k=1}^M \delta_{u^{\Delta,k}(\omega; x, t)}. \quad (117)$$

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

The approximate solutions $v^{\Delta,M}$ were proved to converge to an entropy measure-valued solution of (116) as $(\Delta, M) \rightarrow (0, \infty)$ in [16]. This convergence is in the weak-* topology on Young measures that amounts to convergence of one-point statistical quantities of interest such as the mean, variance, point statistics and probability density functions etc. Thus, the results of figure 4 are justified, mathematically. Furthermore, the computations of [16] also demonstrated that *the measure-valued solution may not be atomic even if the initial data is an atomic young measure concentrated on a L^1 function. Thus, there seems to be no fully deterministic version of multi-dimensional systems of conservation laws and the initial value problem for (115), hitherto considered deterministically, should in fact be considered as a problem in uncertainty quantification (UQ).*

Measure-valued solutions for (116) exist for all times ([19] and references therein) and are able to capture limits of numerical approximation of systems of conservation laws (115). They also serve as an UQ framework within which the random initial data is represented as a Young measure (or point probability distribution). However, measure-valued solutions are not necessarily unique, particularly when the initial

data is non-atomic. This holds true even for scalar conservation laws, see for instance example 9.1 in [19], even for the one-dimensional Burgers' equation. Thus, measure-valued solutions need to be augmented with additional constraints in order to recover uniqueness. This is consistent with the observations, reported in [16, 19], that the computed measure valued solution when realized as a limit of the MC-FV algorithm, is stable with respect to the choice of numerical method and with respect to perturbations of the underlying random initial data. We refer to [40] for numerical approximation of EMV solutions and convergence analysis for a combined MC-FV method for the velocity formulation of the incompressible Euler equations.

7.2 Statistical solutions

An attempt to constrain measure-valued solutions in order to recover uniqueness has been made recently in [17]. In this paper, the authors propose a concept of *statistical solutions* of systems of conservation laws as a suitable solution paradigm as well as computational UQ framework. Statistical solutions of the Navier-Stokes equations in the sense of Foias and Prodi [21] are time-parametrized families of probability measures on a L^p function space. We refer to the surveys [20, 41] for their mathematical theory for the incompressible Navier-Stokes equations.

In these references, statistical solutions are time-dependent probability measures on divergence free L^2 functions that evolve based on either *Cylindrical moments* (Liouville equation) or *Characteristic functionals* (Hopf equation) resulting in a functional differential equation on an infinite-dimensional space.

Although a viable concept for viscous flows such as the incompressible Navier-Stokes equations, it is unclear how the statistical solutions in the sense of [20] can be extended to inviscid problems such as systems of conservation laws (115). Moreover, probability measures on L^p spaces are *non-local* and local statistical quantities such as one-point statistics or multi-point correlations are hard to interpret in this setting. Hence, the linkage between statistical solutions (in the sense of [20] or of the closely related notion of statistical solutions introduced by Vishik and Fursikov in [62]) and measure-valued solutions is unclear.

These difficult issues were tackled in a recent paper [17] in which the authors were able to *localize* probability measures on infinite-dimensional function spaces by relating them to Young measures as described below.

7.2.1 Correlation measures

Let $\mathbf{U} = \mathbb{R}^N$ and $D^k := D \times \dots \times D$ denote the k -fold cartesian product of D . In [17], the authors defined correlation measures as a collection $\mathbf{v} = (v^1, v^2, \dots)$ of maps $v^k : D^k \rightarrow \mathcal{P}(\mathbf{U}^k)$ satisfying the following properties:

- (i.) *Weak-* measurability*: Each map $\mathbf{v}^k : D^k \rightarrow \mathcal{P}(U^k)$ is weak-* measurable, in the sense that the map $x \mapsto \langle \mathbf{v}_x^k, f \rangle$ from $x \in D^k$ into \mathbb{R} is Borel measurable for all $f \in C_0(U^k)$ and $k \in \mathbb{N}$. In other words, \mathbf{v}^k is a Young measure from D^k to U^k .
- (ii.) *L^p -boundedness*: \mathbf{v} is L^p -bounded, in the sense that

$$\int_D \langle \mathbf{v}_x^1, |\xi|^p \rangle dx < +\infty.$$

- (iii.) *Symmetry*: If σ is a permutation of $\{1, \dots, k\}$ and $f \in C_0(\mathbb{R}^k)$ then $\langle \mathbf{v}_{\sigma(x)}^k, f(\sigma(\xi)) \rangle = \langle \mathbf{v}_x^k, f(\xi) \rangle$ for a.e. $x \in D^k$. Here, we denote $\sigma(x) = \sigma(x_1, x_2, \dots, x_k) = (x_{\sigma_1}, x_{\sigma_2}, \dots, x_{\sigma_k})$.
- (iv.) *Consistency*: If $f \in C_0(\mathbf{U}^k)$ is of the form $f(\xi_1, \dots, \xi_k) = g(\xi_1, \dots, \xi_{k-1})$ for some $g \in C_0(\mathbf{U}^{k-1})$, then $\langle \mathbf{v}_{x_1, \dots, x_k}^k, f \rangle = \langle \mathbf{v}_{x_1, \dots, x_{k-1}}^{k-1}, g \rangle$ for almost every $(x_1, \dots, x_k) \in D^k$.
- (v.) *Diagonal continuity (DC)*: If $B_r(x) := \{y \in D : |x - y| < r\}$ then

$$\lim_{r \rightarrow 0} \int_D \frac{1}{|B_r(0)|} \int_{B_r(x)} \langle \mathbf{v}_{x,y}^2, |\xi_1 - \xi_2|^p \rangle dy dx = 0. \quad (118)$$

Each element \mathbf{v}^k is called a *correlation marginal*. The consistency property implies that the k -th correlation marginal \mathbf{v}^k determines all $\mathbf{v}^l, l \leq k$. Thus, the family of correlation marginals is a *hierarchy of young measures*. The *equivalence between correlation measures and probability measures on L^p* is described by the following result, which is [17, Thm. 2.7]:

For every correlation measure \mathbf{v} defined as above, there exists a unique probability measure $\mu \in \mathcal{P}(L^p(D))$ satisfying

$$\int_{L^p} \|u\|_{L^p}^p d\mu(u) < \infty \quad (119)$$

such that, for all $k \in \mathbb{N}$ and for every $g \in L^1(D^k : C_0(\mathbf{U}^k))$: there holds

$$\int_{D^k} \int_{\mathbf{U}^k} g(x, \xi) d\mathbf{v}_x^k(\xi) dx = \int_{L^p} \int_{D^k} g(x, u(x)) dx d\mu(u) \quad (120)$$

(where $u(x)$ denotes the vector $(u(x_1), \dots, u(x_k))$). Conversely, for every probability measure $\mu \in \mathcal{P}(L^p(D))$ with finite L^p bound, there exists a unique correlation measure \mathbf{v} satisfying (120). The relation (120) is also valid for any measurable $g : D \times \mathbf{U} \rightarrow \mathbb{R}$ such that $|g(x, \xi)| \leq C|\xi|^p$ for a.e. $x \in D$.

Moreover, it was also shown in [17] (Theorem 2.20) that the probability measure μ (equivalently the associated correlation measure \mathbf{v}) was uniquely determined in terms of *moments* or *correlation functions* of the correlation measure \mathbf{v} given by

$$m^k : D^k \rightarrow \mathbf{U}^k, \quad m^k(x) := \int_{\mathbf{U}^k} \xi_1 \otimes \dots \otimes \xi_k d\mathbf{v}_x^k(\xi), \quad k \in \mathbb{N}. \quad (121)$$

Here, $\mathbf{U}^{\otimes k}$ refers to the tensor product space $\mathbf{U} \otimes \cdots \otimes \mathbf{U} \simeq \mathbb{R}^{kN}$ (repeated k times), and $\xi_1 \otimes \cdots \otimes \xi_k$ is a functional defined by its action on the dual space $(\mathbf{U}^{\otimes k})^* = \mathbf{U}^{\otimes k}$ through

$$(\xi_1 \otimes \cdots \otimes \xi_k) : (\zeta_1 \otimes \cdots \otimes \zeta_k) = (\xi_1 \cdot \zeta_1) \cdots (\xi_k \cdot \zeta_k).$$

7.2.2 Definition of statistical solutions

Once it is established that probability measures on L^p are completely characterized by moments (correlation functions) of the associated correlation measure, one can evolve an initial probability measure on $L^p(D)$ in time by writing evolution equations for these correlation functions. Following [17], we define statistical solution of (115a) with an initial data $\bar{\mu} \in \mathcal{P}(L^1(\mathbb{R}^d, \mathbb{R}^N))$ as a weak*-measurable mapping $t \mapsto \mu_t \in \mathcal{P}(L^1(\mathbb{R}^d, \mathbb{R}^N))$ such that the corresponding correlation measures $(\nu_t^k)_{k \in \mathbb{N}}$ satisfy the following equations in the sense of distributions,

$$\partial_t \langle \nu_{t,x}^k, \xi_1 \otimes \cdots \otimes \xi_k \rangle + \sum_{i=1}^k \nabla_{x_i} \cdot \langle \nu_{t,x}^k, \xi_1 \otimes \cdots \otimes f(\xi_i) \otimes \cdots \otimes \xi_k \rangle = 0, \quad \forall k \in \mathbb{N}. \quad (122)$$

Note that the first equation in the hierarchy for $k = 1$ precisely agrees with the definition of measure-valued solutions (116). Thus, a *statistical solution is a measure-valued solution that includes information about the evolution of all possible multi-point correlations* in the underlying functions. Hence, statistical solutions are considerably more constrained than measure-valued solutions providing some hope for uniqueness. Moreover, statistical solutions reduce to standard weak solutions as long as the initial data and the resulting statistical solution are *atomic* i.e. $\bar{\mu} = \delta_{\bar{u}}, \mu_t = \delta_{u(t)}$ with $\bar{u}, u(t) \in L^p(D)$. On the other hand, a non-atomic initial probability measure $\bar{\mu}$ can be used to model input uncertainty. Hence, statistical solutions provide a UQ framework [17].

Currently, well-posedness results for statistical solutions are only available for scalar conservation laws. In [17], a concept of entropy statistical solutions was proposed for scalar conservation laws, Definition 4.3 therein. This concept generalizes Kružkhov entropies to probability measures on L^1 . Well-posedness of *entropy statistical solutions* for scalar conservation laws was shown in [17, Thm. 4.7]. These entropy statistical solutions were also shown to satisfy a non-expansive property with respect to the 1-Wasserstein metric on probability measures on L^1 . The mathematical analysis of the convergence of the MC-FV algorithms 1, 2, in the sense of [16], for scalar conservation laws is currently in progress. Some (preliminary) findings are as follows: the same Monte Carlo ensemble averaging algorithm 2, proposed in [16], also converges, under additional assumptions, to a statistical solution of the underlying nonlinear, hyperbolic system as will be shown in a forthcoming paper. Thus, statistical solutions may provide a suitable mathematical and numerical solution framework for multi-dimensional systems of conservation laws as well as of computational uncertainty quantification for them.

The algorithms proposed in [16, 19] are Monte Carlo based. A MLMC version of this algorithm was designed and shown to converge to an entropy measure valued solution in a recent paper [45]. Moreover, it was shown in [45] that if the variance of the details, similar to (96), converge at an algebraic rate i.e if $s > 0$ in (96), then the MLMC algorithm for approximating entropy measure-valued solutions will be more efficient than the Monte Carlo version. However, such an estimate may not hold as shown in figure 4 and the MC and MLMC versions will be comparable (see figure 4 bottom left).

Currently, computation of k -point correlation functions within the framework of statistical solutions uses a full tensor format. This can be prohibitively expensive for even moderate k .

The adaptation of the sparse-tensor algorithms from [47] as described in Section 4.4 to this framework in order to accelerate the computations of multi-point statistical quantities of interest is currently under development.

References

1. Abgrall, R., Harten, A.: Multiresolution representation in unstructured meshes. *SIAM J. Numer. Anal.* **35**(6), 2128–2146 (electronic) (1998). DOI 10.1137/S0036142997315056. URL <http://dx.doi.org/10.1137/S0036142997315056>
2. Barth, A., Schwab, C., Zollinger, N.: Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numer. Math.* **119**(1), 123–161 (2011). DOI 10.1007/s00211-011-0377-0. URL <http://dx.doi.org/10.1007/s00211-011-0377-0>
3. Bierig, C., Chernov, A.: Estimation of arbitrary order central statistical moments by the multilevel Monte Carlo method. *Stoch. Partial Differ. Equ. Anal. Comput.* **4**(1), 3–40 (2016). DOI 10.1007/s40072-015-0063-9. URL <http://dx.doi.org/10.1007/s40072-015-0063-9>
4. Bihari, B.L., Harten, A.: Multiresolution schemes for the numerical solution of 2-D conservation laws. I. *SIAM J. Sci. Comput.* **18**(2), 315–354 (1997). DOI 10.1137/S1064827594278848. URL <http://dx.doi.org/10.1137/S1064827594278848>
5. Bijl, H., Lucor, D., Mishra, S., Schwab, C.: Uncertainty quantification in computational fluid dynamics. In: *Lecture Notes in Computational Science and Engineering 92*. Springer (2013). DOI <http://dx.doi.org/10.1007/978-3-319-00885-1>
6. C. W. Shu and S. Osher: Efficient implementation of essentially non-oscillatory schemes - II. *J. Comput. Phys.* **83**(1), 32–78 (1989)
7. Cockburn, B., Coquel, F., LeFloch, P.G.: Convergence of the finite volume method for multidimensional conservation laws. *SIAM J. Numer. Anal.* **32**(3), 687–705 (1995). DOI 10.1137/0732032. URL <http://dx.doi.org/10.1137/0732032>
8. Cockburn, B., Shu, C.W.: TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comput.* **52**, 411–435 (1989)
9. Cox, S., Hutzenthaler, M., Jentzen, A., van Neerven, J., Welti, T.: Convergence in Hölder norms with applications to Monte Carlo methods in infinite dimensions. Tech. Rep. 2016-28, Seminar for Applied Mathematics, ETH Zürich (2016)
10. Da Prato, G., Zabczyk, J.: Stochastic equations in infinite dimensions, *Encyclopedia of Mathematics and its Applications*, vol. 44. Cambridge University Press, Cambridge (1992). DOI 10.1017/CBO9780511666223. URL <http://dx.doi.org/10.1017/CBO9780511666223>
11. Dafermos, C.M.: Hyperbolic conservation laws in continuum physics, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol.

- 325, third edn. Springer-Verlag, Berlin (2010). DOI 10.1007/978-3-642-04048-1. URL <http://dx.doi.org/10.1007/978-3-642-04048-1>
12. Daun, T., Heinrich, S.: Complexity of Banach space valued and parametric integration. In: Monte Carlo and quasi-Monte Carlo methods 2012, *Springer Proc. Math. Stat.*, vol. 65, pp. 297–316. Springer, Heidelberg (2013). DOI 10.1007/978-3-642-41095-6_12. URL http://dx.doi.org/10.1007/978-3-642-41095-6_12
 13. Daun, T., Heinrich, S.: Complexity of parametric initial value problems in Banach spaces. *J. Complexity* **30**(4), 392–429 (2014). DOI 10.1016/j.jco.2014.01.002. URL <http://dx.doi.org/10.1016/j.jco.2014.01.002>
 14. E. Chiodaroli, C. De Lellis, O. Kreml: Global ill-posedness of the isentropic system of gas dynamics. *Comm. Pure Appl. Math.* **68**, 1157–1190 (2015)
 15. Eymard, R., Gallouët, T., Herbin, R.: Finite volume methods. In: Handbook of numerical analysis, Vol. VII, *Handb. Numer. Anal.*, VII, pp. 713–1020. North-Holland, Amsterdam (2000)
 16. Fjordholm, U., Kappeli, R., Mishra, S., Tadmor, E.: Construction of approximate entropy measure valued solutions for systems of conservation laws. *Journal of Found. Comput. Math* **17**(3), 763–827 (2017). DOI <http://dx.doi.org/10.1007/s10208-015-9299-z>
 17. Fjordholm, U., Lanthaler, S., Mishra, S.: Statistical solutions of hyperbolic conservation laws i: Foundations (2017). Preprint arXiv:1605.05960, *Arch. Rat. Mech. Anal.*, to appear
 18. Fjordholm, U., Mishra, S., Tadmor, E.: Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws. *SIAM J. Numer. Anal.* **50**(2), 544–573 (2012). DOI <http://dx.doi.org/10.1137/110836961>
 19. Fjordholm, U.S., Mishra, S., Tadmor, E.: On the computation of measure-valued solutions. *Acta Numer.* **25**, 567–679 (2016). DOI 10.1017/S0962492916000088. URL <http://dx.doi.org/10.1017/S0962492916000088>
 20. Foias, C., Manley, O., Rosa, R., Temam, R.: Navier-Stokes Equations and Turbulence. Cambridge University Press, Cambridge (2001)
 21. Foias, C., Prodi, G.: Sur les solutions statistiques des équations de Navier-Stokes. *Ann. Mat. Pura Appl.* (4) **111**, 307–330 (1976)
 22. Friz, P.K., Gess, B.: Stochastic scalar conservation laws driven by rough paths. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **33**(4), 933–963 (2016). DOI 10.1016/j.anihpc.2015.01.009. URL <http://dx.doi.org/10.1016/j.anihpc.2015.01.009>
 23. Fuchs, F., McMurry, A., Mishra, S., Risebro, N.H., Waagan, K.: Approximate riemann solver based high-order finite volume schemes for the mhd equations in multi-dimensions. *Communications in Computational Physics* **9**, 324–362 (2011). DOI <http://dx.doi.org/10.4208/cicp.171109.070510a>
 24. Geraci, G., Congedo, P.M., Abgrall, R., Iaccarino, G.: A novel weakly-intrusive non-linear multiresolution framework for uncertainty quantification in hyperbolic partial differential equations. *J. Sci. Comput.* **66**(1), 358–405 (2016). DOI 10.1007/s10915-015-0026-3. URL <http://dx.doi.org/10.1007/s10915-015-0026-3>
 25. Gerhard, N., Müller, S.: Adaptive multiresolution discontinuous Galerkin schemes for conservation laws: multi-dimensional case. *Comput. Appl. Math.* **35**(2), 321–349 (2016). DOI 10.1007/s40314-014-0134-y. URL <http://dx.doi.org/10.1007/s40314-014-0134-y>
 26. Gess, B., Perthame, B., Souganidis, P.E.: Semi-discretization for stochastic scalar conservation laws with multiple rough fluxes. *SIAM J. Numer. Anal.* **54**(4), 2187–2209 (2016). DOI 10.1137/15M1053670. URL <http://dx.doi.org/10.1137/15M1053670>
 27. Giles, M.B.: Multilevel Monte Carlo methods. *Acta Numer.* **24**, 259–328 (2015). DOI 10.1017/S096249291500001X. URL <http://dx.doi.org/10.1017/S096249291500001X>
 28. Godlewski, E., Raviart, P.A.: Hyperbolic systems of conservation laws, *Mathématiques & Applications (Paris) [Mathematics and Applications]*, vol. 3/4. Ellipses, Paris (1991)
 29. Godlewski, E., Raviart, P.A.: Numerical approximation of hyperbolic systems of conservation laws, *Applied Mathematical Sciences*, vol. 118. Springer-Verlag, New York (1996)
 30. Gottlieb, S., Shu, C., Tadmor, E.: High order time discretizations with strong stability property. *SIAM Review* **43**, 89–112 (2001)

31. Griebel, M., Harbrecht, H.: A note on the construction of L -fold sparse tensor product spaces. *Constr. Approx.* **38**(2), 235–251 (2013). DOI 10.1007/s00365-012-9178-7. URL <http://dx.doi.org/10.1007/s00365-012-9178-7>
32. Harbrecht, H., Peters, M., Siebenmorgen, M.: Combination technique based k -th moment analysis of elliptic problems with random diffusion. *J. Comput. Phys.* **252**, 128–141 (2013). DOI 10.1016/j.jcp.2013.06.013. URL <http://dx.doi.org/10.1016/j.jcp.2013.06.013>
33. Harten, A., Engquist, B., Osher, S., Chakravarty, S.R.: Uniformly high order accurate essentially non-oscillatory schemes. *J. Comput. Phys.* pp. 231–303 (1987)
34. Hildebrand, A., Mishra, S.: Entropy stable shock capturing streamline diffusion space-time discontinuous Galerkin (DG) methods for systems of conservation laws. *Numer. Math.* **126**(1), 103–151 (2014)
35. Holden, H., Risebro, N.H.: Front tracking for hyperbolic conservation laws, *Applied Mathematical Sciences*, vol. 152. Springer, New York (2011). DOI 10.1007/978-3-642-23911-3. URL <http://dx.doi.org/10.1007/978-3-642-23911-3>. First softcover corrected printing of the 2002 original
36. Koley, U., Risebro, N.H., Schwab, C., Weber, F.: A multilevel monte carlo finite difference method for random scalar degenerate convection diffusion equations. *Journal of Hyperbolic Differential Equations* (2017). (to appear).
37. Kröner, D.: Numerical schemes for conservation laws. Wiley-Teubner Series Advances in Numerical Mathematics. John Wiley & Sons Ltd., Chichester (1997)
38. Ledoux, M., Talagrand, M.: Probability in Banach spaces. *Classics in Mathematics*. Springer-Verlag, Berlin (2011). Isoperimetry and processes, Reprint of the 1991 edition
39. LeFloch, P.G.: Hyperbolic systems of conservation laws. *Lectures in Mathematics ETH Zürich*. Birkhäuser Verlag, Basel (2002). DOI 10.1007/978-3-0348-8150-0. URL <http://dx.doi.org/10.1007/978-3-0348-8150-0>. The theory of classical and nonclassical shock waves
40. Leonardi, F.: A projection method for the computation of admissible measure valued solutions of the incompressible euler equations. Tech. Rep. 2017-06, Seminar for Applied Mathematics, ETH Zürich (2017)
41. Leonardi, F., Mishra, S., Schwab, C.: Numerical approximation of statistical solutions of planar, incompressible flows. *Math. Models and Method in Applied Sciences* **26**(13), 2471–2524 (2016). DOI <http://dx.doi.org/10.1142/S0218202516500597>
42. LeVeque, R.J.: Finite volume methods for hyperbolic problems. *Cambridge Texts in Applied Mathematics*. Cambridge University Press, Cambridge (2002). DOI 10.1017/CBO9780511791253. URL <http://dx.doi.org/10.1017/CBO9780511791253>
43. Lions, P.L., Perthame, B., Souganidis, P.E.: Scalar conservation laws with rough (stochastic) fluxes. *Stoch. Partial Differ. Equ. Anal. Comput.* **1**(4), 664–686 (2013)
44. Lions, P.L., Perthame, B., Souganidis, P.E.: Scalar conservation laws with rough (stochastic) fluxes: the spatially dependent case. *Stoch. Partial Differ. Equ. Anal. Comput.* **2**(4), 517–538 (2014). DOI 10.1007/s40072-014-0038-2. URL <http://dx.doi.org/10.1007/s40072-014-0038-2>
45. Lye, K.: Multilevel monte-carlo for measure valued solutions. Tech. Rep. 2016-51 (revised), Seminar for Applied Mathematics, ETH Zürich (2016)
46. Mishra, S., Risebro, N.H., Schwab, C., Tokareva, S.: Numerical solution of scalar conservation laws with random flux functions. *SIAM/ASA J. Uncertainty Quantification* **4**(1), 552–591 (2016). DOI <http://dx.doi.org/10.1137/120896967>
47. Mishra, S., Schwab, C.: Sparse Tensor Multi-Level Monte Carlo Finite Volume Methods for hyperbolic conservation laws with random initial data. *Mathematics of Computation* **81**, 1979–2018 (2012). DOI <http://dx.doi.org/10.1090/S0025-5718-2012-02574-9>
48. Mishra, S., Schwab, C., Sukys, J.: Multi-level monte carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions. *Journal of Computational Physics* **231**(8), 3365–3388 (2012). DOI <http://dx.doi.org/10.1016/j.jcp.2012.01.011>
49. Mishra, S., Schwab, C., Sukys, J.: Multi-level Monte Carlo finite volume methods for uncertainty quantification of acoustic wave propagation in random heterogeneous layered medium. *J. Comput. Phys.* **312**, 192–217 (2016). DOI 10.1016/j.jcp.2016.02.014. URL <http://dx.doi.org/10.1016/j.jcp.2016.02.014>

50. Müller, S.: Multiresolution schemes for conservation laws. In: Multiscale, nonlinear and adaptive approximation, pp. 379–408. Springer, Berlin (2009). DOI 10.1007/978-3-642-03413-8_11. URL http://dx.doi.org/10.1007/978-3-642-03413-8_11
51. van Neerven, J.: Stochastic evolution equations (2007/8). Lecture Notes, ISEM
52. von Petersdorff, T., Schwab, C.: Sparse finite element methods for operator equations with stochastic data. *Appl. Math.* **51**(2), 145–180 (2006). DOI 10.1007/s10492-006-0010-1. URL <http://dx.doi.org/10.1007/s10492-006-0010-1>
53. R. J. DiPerna: Measure-valued solutions to conservation laws. *Archive for Rational Mechanics and Analysis* **88**, 223–270 (1985)
54. Risebro, N.H., Schwab, C., Weber, F.: Multilevel monte carlo front-tracking for random scalar conservation laws. *BIT Numerical Mathematics* (2016). DOI <http://dx.doi.org/10.1007/s10543-015-0550-4>
55. Schmidlin, G., Schwab, C.: Wavelet Galerkin BEM on unstructured meshes by aggregation. In: Multiscale and multiresolution methods, *Lect. Notes Comput. Sci. Eng.*, vol. 20, pp. 359–378. Springer, Berlin (2002). DOI 10.1007/978-3-642-56205-1_12. URL http://dx.doi.org/10.1007/978-3-642-56205-1_12
56. Schwab, C., Gittelsohn, C.J.: Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numer.* **20**, 291–467 (2011). DOI 10.1017/S0962492911000055. URL <http://dx.doi.org/10.1017/S0962492911000055>
57. Schwab, C., Tokareva, S.: High order approximation of probabilistic shock profiles in hyperbolic conservation laws with uncertain initial data. *ESAIM: Mathematical Modelling and Numerical Analysis* **47**(3), 807–835 (2013). DOI <http://dx.doi.org/10.1051/m2an/2012060>
58. Smoller, J.: Shock waves and reaction-diffusion equations, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 258, second edn. Springer-Verlag, New York (1994)
59. Sukys, J.: Robust multi-level monte carlo finite volume methods for systems of conservation laws with random input data. Ph.D. thesis, ETH Zürich (2014)
60. Todor, R.A.: A new approach to energy-based sparse finite-element spaces. *IMA J. Numer. Anal.* **29**(1), 72–85 (2009). DOI 10.1093/imanum/drm041. URL <http://dx.doi.org/10.1093/imanum/drm041>
61. Tryoen, J., Le Maître, O., Ndjinga, M., Ern, A.: Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems. *J. Comput. Phys.* **229**(18), 6485–6511 (2010). DOI 10.1016/j.jcp.2010.05.007. URL <http://dx.doi.org/10.1016/j.jcp.2010.05.007>
62. Višik, M.I., Fursikov, A.V.: Solutions statistiques homogènes des systèmes différentiels paraboliques et du système de Navier-Stokes. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **4**(3), 531–576 (1977)
63. Wehr, J., Xin, J.: Front speed in the Burgers equation with a random flux. *J. Statist. Phys.* **88**(3-4), 843–871 (1997). DOI 10.1023/B:JOSS.0000015175.70862.77. URL <http://dx.doi.org/10.1023/B:JOSS.0000015175.70862.77>