Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random intitial data

S. Mishra and Ch. Schwab

Research Report No. 2010-24
September 2010

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

.

SPARSE TENSOR MULTI-LEVEL MONTE CARLO FINITE VOLUME METHODS FOR HYPERBOLIC CONSERVATION LAWS WITH RANDOM INITIAL DATA

S. MISHRA AND CH. SCHWAB

Abstract. We consider scalar hyperbolic conservation laws in several ($d \geq 1$) spatial dimensions with stochastic initial data. We prove existence and uniqueness of a random-entropy solution and show existence of statistical moments of any order k of this random entropy solution. We present a class of numerical schemes of multi-level Monte Carlo Finite Volume (MLMC-FVM) type for the approximation of random entropy solutions as well as of their k-point correlation functions. These schemes are shown to obey the same accuracy vs. work estimate as a single application of the finite volume solver for the corresponding deterministic problem. Numerical experiments demonstrating the efficiency of these schemes are presented. Statistical moments of discontinuous solutions are found to be more regular than pathwise solutions.

1. Introduction

Many problems in physics and engineering are modeled by hyperbolic systems of conservation laws. The Cauchy problem for such systems takes the form

$$\begin{aligned} &\mathbf{U}_t + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}(\mathbf{F}_j(\mathbf{U})) = 0, \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d, t > 0, \\ &\mathbf{U}(x,0) = \mathbf{U}_0(x), \quad x \in \mathbb{R}^d. \end{aligned} \tag{1.1}$$

Here, $\mathbf{U} : \mathbb{R}^d \mapsto \mathbb{R}^m$ is the vector of unknowns and $\mathbf{F}_j : \mathbb{R}^m \mapsto \mathbb{R}^m$ is the flux vector for the j-th direction with m being a positive integer.

Examples include the Shallow Water Equations of hydrology, the Euler Equations for inviscid, compressible flow and the Magnetohydrodynamic (MHD) equations of plasma physics, see, e.g. [5, 9]. An illustrative model for (1.1) is provided by taking $m = 1$ and obtaining the so-called scalar conservation law.

It is well known that solutions of (1.1) develop discontinuities in finite time even when the initial data is smooth. This holds true even for the scalar case and solutions to (1.1) are sought in the weak sense. Furthermore, weak solutions are augmented by additional admissibility criteria or *entropy conditions* ([5]) in order to ensure uniqueness. Well-posedness of entropy solutions in the scalar case (even for several space dimensions) was obtained by Kruzkhov. Some local well-posedness results for systems in one space dimension exist but no global well-posedness results for systems of conservation laws are available in several space dimensions.

Date: September 1, 2010.

1991 *Mathematics Subject Classification.* 65N30,65M06,35L65.

Acknowledgement. The work of CS was supported in part by ERC grant no. 247277. CS and SM acknowledge also partial support from ETH grant no. CH1-03 10-1. SM wishes to thank Claude J. Gittelson for useful discussions.

Numerical methods for approximating entropy solutions of systems of conservation laws have undergone extensive development and many efficient methods are available, see [9, 17]. In particular, finite volume methods are frequently employed for approximating (1.1).

This *classical* paradigm for designing efficient numerical schemes assumes that the initial data $\mathbf{U}_0$ in (1.1) is known exactly. However, in many situations of practical interest, it is not possible to obtain the initial data exactly due to inherent uncertainty in measurements. Then, the initial data are known only upto certain statistical quantities of interest like the mean, variance, higher moments and in some cases, the law of the stochastic initial data; in such cases, a mathematical formulation of (1.1) is required which allows for *random initial data*.

The *first* aim of this paper is to develop an appropriate mathematical framework of *random* entropy solutions for conservation laws. As the theory for deterministic initial data is only well-developed in the scalar case, we focus here on a scalar conservation law in spatial dimension $d \geq 1$ with random initial data. We define random entropy solutions and provide an existence and uniqueness result, generalizing the classical well-posedness results of Kruzkhov to the case of uncertain initial data. Furthermore, we prove the existence of statistical quantities of the random entropy solution like the statistical mean and two and k-point spatial and temporal correlation functions under suitable regularity assumptions on the initial data. In particular, we show that if the initial data has finite statistical moments of order k in $L^1(\mathbb{R}^d)$, the random entropy solution also possesses finite moments of order k as well, for any k. We remark that randomness in the initial data is just one of the many available mechanisms for introducing uncertainty in the solutions of (1.1). One could consider also random boundary data (if we consider (1.1) in a bounded domain $D \subset \mathbb{R}^d$) or random source terms added to (1.1). Let us mention that hyperbolic conservation laws with various types of random data were considered in the literature (see, e.g., [14, 15, 6] and the references there).

The *second* aim of this paper is to design *fast* and robust numerical algorithms for computing random entropy solutions. In particular, we focus on statistical sampling techniques of the Monte Carlo (MC) type. MC methods consist of *sampling* the probability space and solving the deterministic version of the underlying PDE for each sample. As we will show, MC methods are "non-intrusive", very easy to code and to parallelize and well suited for random solutions with low spatial regularity as in conservation laws where discontinuities are generic. However,as we shall prove, MC methods converge only at rate $1/2$ as the number M of MC samples increases thereby requiring a large number of samples in order to obtain low statistical errors. This slow convergence entails high computational costs for MC type methods.

In order to deal with the aforementioned issues, we device a novel *Multi-level Monte Carlo* (MLMC) algorithm based on finite volume schemes for the deterministic version of the conservation law. This family of methods was introduced by S. Heinrich for numerical quadrature ([12]) and by M. Giles in the context of path simulations for stochastic ordinary differential equations ([7, 8]). More recently, MLMC finite element methods for elliptic problems with stochastic coefficients were introduced by Barth, Schwab and Zollinger in [2].

In the current paper, the first analysis of the MLMC finite volume method (MLMC-FVM) is presented in the context of a scalar conservation law in several

space dimensions. In particular, the MLMC-FVM is shown to converge. Furthermore, our convergence analysis yields an optimal strategy for chosing MC samples depending on the spatial and temporal meshwidth. This allows us to prove for MLMC-FVM an accuracy vs. work estimate that equals, in two and three spatial dimensions and for first and second order schemes, the corresponding estimate *for the FVM solution of a single, deterministic problem* (1.1). This contrasts very sharply with the accuracy vs. work estimate of the single level MC method. In particular, our mathematical convergence analysis and the extensive numerical experiments in the present paper show that the MLMC-FVM algorithm provides a *fast*, robust, non-intrusive and highly parallelizable recipe for computing random entropy solutions of conservation laws with uncertainty. We also introduce a sparse tensor formalism that allows us to approximate higher statistical moments of the random entropy solutions of (1.1) with (up to logarithmic factors of the meshwidth and the time step) the same error vs. work as a single, deterministic FVM solve for (1.1).

At this juncture, we would like to remark that statistical MC type methods for random conservation laws have not been as widely studied as stochastic Galerkin methods based on generalized Polynomial Chaos (gPC for short). An incomplete list of references for gPC methods for uncertainty quantification in hyperbolic conservation laws includes [1, 3, 19, 26, 22, 27] and other references therein. Despite some advantages, these gPC methods are more intrusive, harder to implement and more difficult to parallelize than MC methods. Hence, in the present paper, we focus on the design, the mathematical analysis and the numerical implementation of MLMC-FVM methods.

The remainder of this paper is organized as follows: in Section 2, we introduce some preliminary notions from probability theory and functional analysis. The concept of random entropy solutions is introduced and shown to be well-posed in Section 3. Numerical schemes are designed in Section 4 and numerical experiments are presented in Section 5.

2. Preliminaries from Probability

Our mathematical formulation of scalar conservation laws with random data will use the concept of random variables taking values in function spaces. We recapitulate basic concepts from Chapter 1 of [4]. Let $(\Omega, \mathcal{F})$ be a measurable space, with Ω denoting the set of all elementary events, and $\mathcal{F}$ a σ-algebra of all possible events in our probability model. If $(E, \mathcal{G})$ denotes a second measurable space, then an *E-valued random variable* (or random variable taking values in E) is any mapping $X : \Omega \to E$ such that the set $\{\omega \in \Omega\colon\ X(\omega) \in A\} = \{X \in A\} \in \mathcal{F}$ for any $A \in \mathcal{G}$, i.e. such that X is a $\mathcal{G}$-measurable mapping from Ω into E.

Assume now that E is a metric space; with the Borel σ-field $\mathcal{B}(E)$, $(E, \mathcal{B}(E))$ is a measurable space and we shall always assume that E-valued random variables $X : \Omega \to E$ will be $(\mathcal{F}, \mathcal{B}(E))$ measurable. If E is a separable Banach-space with norm $\|\circ\|_E$ and (topological) dual E^*, then $\mathcal{B}(E)$ is the smallest σ-field of subsets of E containing all sets

$$\{x \in E : \varphi(x) \leq \alpha\},\ \varphi \in E^*,\ \alpha \in \mathbb{R}\,. \tag{2.1}$$

Hence if E is a separable Banach space, $X : \Omega \to E$ is an E-valued random variable iff for every $\varphi \in E^*$, $\omega \longmapsto \varphi(X(\omega)) \in \mathbb{R}^1$ is an $\mathbb{R}^1$-valued random variable. Moreover, we have

Lemma 2.1. *Let E be a separable Banach-space and let $X : \Omega \to E$ be an E-valued random variable on $(\Omega, \mathcal{F})$. Then the mapping $\Omega \ni \omega \longmapsto \|X(\omega)\|_E \in \mathbb{R}^1$ is measurable.*

Proof. Since E is separable, there exists a sequence $\{\varphi_n\} \subset E^*$ such that for all $x \in E$ holds

$$\|x\|_E = \sup_{n\in\mathbb{N}} |\varphi_n(x)| \,. \tag{2.2}$$

Hence we find

$$\forall \omega \in \Omega : \ \|X(\omega)\|_E = \sup_{n\in\mathbb{N}} |\varphi_n(X(\omega))| \tag{2.3}$$

which implies that $\omega \longmapsto \|X(\omega)\|_E$ is an $\mathbb{R}^1$-valued random variable. □

The random variable $X : \Omega \to E$ is called *Bochner integrable* if, for any probability measure $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{F})$,

$$\int_\Omega \|X(\omega)\|_E \, \mathbb{P}(d\omega) < \infty \,. \tag{2.4}$$

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

If $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ is a random variable, $\mathcal{L}(X)$ denotes the *law of X under* $\mathbb{P}$, i.e.

$$\mathcal{L}(X)(A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}) \quad \forall A \in \mathcal{E} \,. \tag{2.5}$$

The image measure $\mu_X = \mathcal{L}(X)$ on $(E, \mathcal{E})$ is called law or distribution of X.

A random variable taking values in E is called *simple* if it can take only finitely many values, i.e. if it has the explicit form (with χ_A the indicator function of $A \in \mathcal{F}$)

$$X = \sum_{i=1}^N x_i \, \chi_{A_i}, \quad A_i \in \mathcal{F}, \ x_i \in E, \ N < \infty \,. \tag{2.6}$$

We set, for simple random variables X taking values in E and for any $\mathcal{B} \in \mathcal{F}$,

$$\int_B X(\omega) \, \mathbb{P}(d\omega) = \int_B X d\,\mathbb{P} := \sum_{i=1}^N x_i \, \mathbb{P}(A_i \cap B) \,. \tag{2.7}$$

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

$$\Big\| \int_B X(\omega) \, \mathbb{P}(d\omega) \Big\|_E \le \int_B \|X(\omega)\|_E \, \mathbb{P}(d\omega) \,. \tag{2.8}$$

For any random variable $X : \Omega \to E$ which is Bochner integrable, there exists a sequence $\{X_m\}_{m\in\mathbb{N}}$ of simple random variables such that, for all $\omega \in \Omega$, $\|X(\omega) -$

$X_m(\omega)\|_E \to 0$ as $m \to \infty$. Therefore, (2.7) and (2.8) extend in the usual fashion by continuity to any E-valued random variable. We denote the integral

$$\int_\Omega X(\omega)\,\mathbb{P}(d\omega) = \lim_{m\to\infty} \int_\Omega X_m(\omega)\,\mathbb{P}(d\omega) \in E \tag{2.9}$$

by $\mathbb{E}[X]$ ("expectation" of X). □

We shall use that operators that act on random variables to generate random variables. The following result is obtained by approximating with simple random variables.

Lemma 2.2. *Assume that E, F are separable Banach spaces, and that $\mathcal{E} = \mathcal{B}(E)$, $\mathcal{F} = \mathcal{B}(F)$. Assume further that $A : D(A) \subset E \to F$ is a closed operator such that the domain $D(A)$ of A satisfies $D(A) \in \mathcal{B}(E)$. If $X : \Omega \to E$ is a random variable such that $X(\omega) \in D(A)$ $\mathbb{P}$-a.s., then it holds $\mathbb{P}$-a.s.*

$$\left|\begin{array}{l} X \text{ is a } D(A)\text{-valued random variable, } \mathbb{P}\text{-a.s., and} \\ AX \text{ is an } F\text{-valued random variable, } \mathbb{P}\text{-a.s.} \end{array}\right. \tag{2.10}$$

Moreover, if

$$\int_\Omega \|AX(\omega)\|_F\,\mathbb{P}(d\omega) < \infty\,, \tag{2.11}$$

then it holds in F that

$$A\int_\Omega X(\omega)\,\mathbb{P}(d\omega) = \int_\Omega AX(\omega)\,\mathbb{P}(d\omega)\,. \tag{2.12}$$

We shall require for $1 \le p \le \infty$ Bochner spaces of p-summable random variables X taking values in the Banach-space E. By $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ we denote the set of all (equivalence classes of) integrable, E-valued random variables X. We equip it with the norm

$$\|X\|_{L^1(\Omega;E)} := \mathbb{E}(\|X\|_E) = \int_\Omega \|X(\omega)\|_E\,\mathbb{P}(d\omega)\,. \tag{2.13}$$

More generally, for $1 \le p < \infty$, we define $L^p(\Omega, \mathcal{F}, \mathbb{P}; E)$ as the set of p-summable random variables taking values E and equip it with norm

$$\|X\|_{L^p(\Omega;E)} := (\mathbb{E}(\|X\|_E^p))^{1/p},\ 1 \le p < \infty\,. \tag{2.14}$$

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

$$\|X\|_{L^\infty(\Omega;E)} := \operatorname{ess\,sup}_{\omega\in\Omega} \|X(\omega)\|_E\,. \tag{2.15}$$

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 $r \ge p \ge 1$,

$$L^r(0, T; E),\ C^0([0, T]; E) \in \mathcal{B}(L^p(0, T; E))\,. \tag{2.16}$$

3. Hyperbolic Conservation Laws with random data

3.1. **Scalar hyperbolic conservation laws.** We consider the Cauchy problem for scalar conservation laws (SCL) by setting $m = 1$ in (1.1) and obtaining,

$$\frac{\partial u}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} \left(f_j(u)\right) = 0, \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d,\ t > 0\,. \tag{3.1}$$

Here the unknown is $u : \mathbb{R}^d \mapsto \mathbb{R}$. Introducing the flux function $f(u)$

$$f(u) = (f_1(u), \ldots, f_d(u)) \in C^1(\mathbb{R}, \mathbb{R}^d)\,, \quad \operatorname{div} f(u) = \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f_j(u)\,, \tag{3.2}$$

we may rewrite (3.1) succinctly as

$$\frac{\partial u}{\partial t} + \operatorname{div}\left(f(u)\right) = 0 \ \text{ in } \ \mathbb{R}^d \times \mathbb{R}_+\,. \tag{3.3}$$

We supply the SCL (3.3) with initial condition

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}^d\,. \tag{3.4}$$

3.2. **Entropy Solution.** The Cauchy problem (3.3), (3.4) admits, for each $u_0 \in L^\infty(\mathbb{R}^d)$, a unique entropy solution (see, e.g., [9, 24, 5]). Moreover, for every $t > 0$, $u(\cdot, t) \in L^1(\mathbb{R}^d)$ and the (nonlinear) data-to-solution operator

$$S : u_0 \longmapsto u(\cdot, t) = S(t)\, u_0, \quad t > 0 \tag{3.5}$$

has several properties which will be crucial for our subsequent development. To state the properties of $\{S(t)\}_{t\geq 0}$, we introduce some additional notation: for a Banach-space E with norm $\|\circ\|_E$, and for $0 < T \leq +\infty$, denote by $C_b(0, T; E)$ the space of bounded and continuous functions from $[0, T]$ into E, and by $L^p(0, T; E)$, $1 \leq p \leq +\infty$, the space of strongly measurable functions from $(0, T)$ to E such that for $1 \leq p < +\infty$

$$\|v\|_{L^p(0,T;E)} = \Big(\int_0^T \|v(t)\|_E^p \, dt \Big)^{\frac{1}{p}}, \tag{3.6}$$

respectively, if $p = \infty$,

$$\|v\|_{L^\infty(0,T;E)} = \operatorname{ess} \sup_{0\leq t\leq T} \|v(t)\|_E \tag{3.7}$$

are finite. The following result summarizes the classical results on existence and uniqueness of an entropy weak solution of the SCL (3.1)-(3.4) (we refer to, eg., [9, 10, 16, 11, 17]).

Theorem 3.1.

1) *For every $u_0 \in L^\infty(\mathbb{R}^d)$, (3.1) - (3.4) admits a unique entropy solution $u \in L^\infty(\mathbb{R}^d \times (0, T)) := L^\infty(0, T; L^\infty(\mathbb{R}^d))$.*
2) *For every $t > 0$, the (nonlinear) data-to-solution map $S(t)$ given by*

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

satisfies

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

$$\|S(t)u_0 - S(t)v_0\|_{L^1(\mathbb{R}^d)} \leq \|u_0 - v_0\|_{L^1(\mathbb{R}^d)}\,. \tag{3.8}$$

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

$$TV(S(t)u_0) \leq TV(u_0) \quad \forall u_0 \in (L^1 \cap BV)(\mathbb{R}^d)\,. \tag{3.9}$$

iii) *For every $u_0 \in (L^\infty \cap L^1)(\mathbb{R}^d)$,*

$$\|S(t)u_0\|_{L^\infty(\mathbb{R}^d)} \leq \|u_0\|_{L^\infty(\mathbb{R}^d)}\,; \tag{3.10}$$

$$\|S(t)u_0\|_{L^1(\mathbb{R}^d)} \leq \|u_0\|_{L^1(\mathbb{R}^d)}\,. \tag{3.11}$$

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

$$\|S(\cdot)u_0\|_{C(0,T;L^1(\mathbb{R}^d))} = \max_{0\leq t\leq T}\|S(t)u_0\|_{L^1(\mathbb{R}^d)} \leq \|u_0\|_{L^1(\mathbb{R}^d)}\,. \tag{3.12}$$

3.3. **Random Initial Data and Random Entropy Solution.** Based on Theorem 3.1, we will now formulate (3.1) - (3.4) for random initial data. To this end, we denote $(\Omega,\mathcal{F},\mathbb{P})$ a probability space, and assume we are given as u_0 a $L^1(\mathbb{R}^d)$-valued random variable, i.e. a $L^1(\mathbb{R}^d)$ measurable map

$$u_0 : (\Omega,\mathcal{F}) \longmapsto \big(L^1(\mathbb{R}^d),\ \mathcal{B}(L^1(\mathbb{R}^d))\big)\,. \tag{3.13}$$

We assume further that

$$u_0(\cdot,\omega) \in L^\infty(\mathbb{R}^d) \cap BV(\mathbb{R}^d) \quad \mathbb{P}\text{-a.s.}, \tag{3.14}$$

which is to say that

$$\mathbb{P}(\{\omega\in\Omega : u_0(\cdot,\omega) \in (L^\infty\cap BV)(\mathbb{R}^d)\}) = 1\,. \tag{3.15}$$

Since $L^1(\mathbb{R}^d)$ is separable, (3.13) is well defined and we may impose for $k\in\mathbb{N}$ the *k-th moment condition*

$$\|u_0\|_{L^k(\Omega;L^1(\mathbb{R}^d))} < \infty\,, \tag{3.16}$$

where the Bochner spaces with respect to the probability measure are defined in Section 2. With these preliminaries in hand, we have the following definition,

Definition 3.2. *A random field $u : \Omega \ni \omega \to u(x,t;\omega)$, i.e. a measurable mapping from $(\Omega,\mathcal{F})$ to $C([0,T];L^1(\mathbb{R}^d))$, is said to be a* random entropy solution *if it satisfies the following,*

(i.) Weak solution: For $\mathbb{P}$-a.e $\omega\in\Omega$, u satisfies the following integral identity,

$$\int_0^\infty\int_{\mathbb{R}^d}\left(u(x,t,\omega)\varphi_t(x,t) + \sum_{j=1}^d f_j(u(x,t,\omega))\frac{\partial}{\partial x_j}\varphi(x,t)\right)dxdt + \int_{\mathbb{R}^d} u_0(x,\omega))\varphi(x,0)dx = 0, \tag{3.17}$$

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

(ii.) Entropy condition: For any entropy-entropy flux pair i.e η, Q_j with $j = 1,2,\dots,d$ are smooth functions such that η is convex and $Q_j' = \eta' f_j'$ for all j, and for $\mathbb{P}$-a.e $\omega\in\Omega$, u satisfies the following integral identity,

$$\int_0^\infty\int_{\mathbb{R}^d}\left(\eta(u(x,t,\omega))\varphi_t(x,t) + \sum_{j=1}^d Q_j(u(x,t,\omega))\frac{\partial}{\partial x_j}\varphi(x,t)\right)dxdt \geq 0, \tag{3.18}$$

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

We then have from Theorem 3.1 the following result:

Theorem 3.3. *Consider the scalar conservation laws* (3.1) - (3.4) *with random initial data* $u_0 : \Omega \to L^1(\mathbb{R}^d)$ *satisfying* (3.14), (3.15) *and the* k*-th moment condition* (3.16) *for some* $k \in \mathbb{N}$. *Then there exists a unique random entropy solution* $u : \Omega \ni \omega \to C_b(0,T;L^1(\mathbb{R}^d))$ *given by*

$$u(\cdot,t;\omega) = S(t)u_0(\cdot,\omega)\,, \quad t>0,\ \omega\in\Omega \tag{3.19}$$

such that for every $k \geq m \geq 1$ *and for every* $0 \leq t \leq T < \infty$ *holds* $\mathbb{P}$*-a.s.*

$$\|u\|_{L^k(\Omega;C(0,T;L^1(\mathbb{R}^d)))} \leq \|u_0\|_{L^k(\Omega;L^1(\mathbb{R}^d))}\,, \tag{3.20}$$

$$\|S(t)\,u_0(\cdot,\omega)\|_{(L^1\cap L^\infty)(\mathbb{R}^d)} \leq \|u_0(\cdot,\omega)\|_{(L^1\cap L^\infty)(\mathbb{R}^d)} \tag{3.21}$$

and such that we have

$$TV(S(t)u_0(\cdot,\omega)) \leq TV(u_0(\cdot,\omega)) \tag{3.22}$$

$\mathbb{P}$*-a.s..*

Proof.

i) For $\omega \in \Omega$, we define, motivated by Theorem 3.1, for $\mathbb{P}$-a.e. $\omega \in \Omega$ a random function $u(t,x;\omega)$ by

$$u(\cdot;\omega) = S(t)\,u_0(\cdot,\omega)\,. \tag{3.23}$$

By the properties of the solution mapping $(S(t))_{t\geq 0}$, the random field defined in (3.23) is well defined; for $\mathbb{P}$-a.e. $\omega \in \Omega$, $u(\cdot;\omega)$ is a weak entropy solution of the SCL (3.1)-(3.4).

ii) From Theorem 3.1, we obtain that $\mathbb{P}$-a.s., all bounds (3.9) - (3.12) hold. We proceed to check measurability of the mapping $\Omega \ni \omega \to u(\cdot\,;\omega) = S(t)u_0(\cdot\,;\omega)$.

iii) We have to show that for every $0 \leq t \leq T$, the mapping $\Omega \ni \omega \longmapsto u(\cdot,t;\omega) = S(t)\,u_0(\cdot;\omega)$ is measurable, as an $L^1(\mathbb{R}^d)$-valued random variable. This, however, is a consequence of the fact that $L^1(\mathbb{R}^d)$ is separable and that $L^1(\mathbb{R}^d)^* \cong LL^\infty(\mathbb{R}^d)$. Thus $\mathcal{B}(L^1(\mathbb{R}^d))$ is the smallest σ-algebra containing all subsets of $L^1(\mathbb{R}^d)$ of the form

$$\{v \in L^1(\mathbb{R}^d) : \varphi(v) \leq \alpha\} : \quad \varphi \in L^\infty(\mathbb{R}^d),\ \alpha \in \mathbb{R}\,.$$

Let now $\alpha \in \mathbb{R}$, $\varphi \in L^\infty(\mathbb{R}^d) = L^1(\mathbb{R}^d)^*$ and $t > 0$. Consider then

$$\{u(\cdot,t;\omega)\} :\ \varphi(u(\cdot,t;\omega)) \leq \alpha\} = \{S(t)u_0(\cdot;\omega) :\ \varphi(S(t)u_0(\cdot;\omega)) \leq \alpha\}\,. \tag{3.24}$$

By continuity (3.11) of $S(t)$ in $L^1(\mathbb{R}^d)$, for every $0 < t \leq T < \infty$, $S(t)$ maps open balls of $L^1(\mathbb{R}^d)$ into open balls of $L^1(\mathbb{R}^d)$. Since $\mathcal{B}(L^1(\mathbb{R}^d))$ is the smallest σ-field on $L^1(\mathbb{R}^d)$ containing all subsets of the form (3.24), and since $u_0 \in L^0(\Omega,\mathcal{F},\mathbb{P};L^1(\mathbb{R}^d))$, we have for every $0 < t \leq T < \infty$ that

$$u(\cdot,t;\cdot) = S(t)u_0(\cdot;\cdot) \in L^0(\Omega,\mathcal{F},\mathbb{P};L^1(\mathbb{R}^d))\,.$$

iv) To show (3.20), assume $u_0 \in L^k(\Omega,\mathcal{F},\mathbb{P};L^1(\mathbb{R}^d))$ for some $k \in \mathbb{N}$. Then, for every $0 \leq t \leq T < \infty$, we have

$$\begin{aligned}\int_\Omega \|u(\cdot,t;\omega)\|^k_{L^1(\mathbb{R}^d)}\mathbb{P}(d\omega) &= \int_\Omega \|S(t)u_0(\cdot;\omega)\|^k_{L^1(\mathbb{R}^d)}\,\mathbb{P}(d\omega)\\ &\overset{(3.11)}{\leq} \int_\Omega \|u_0(\cdot;\omega)\|^k_{L^1(\mathbb{R}^d)}\mathbb{P}(d\omega) = \|u_0\|^k_{L^k(\Omega,\mathcal{F},\mathbb{P};L^1(\mathbb{R}^d))}\,.\end{aligned}$$

Hence

$$\begin{aligned}\|u_0\|^k_{L^k(\Omega,\mathcal{F},\mathbb{P};C(0,T;L^1(\mathbb{R}^d)))} &= \int_\Omega \big(\max_{0\le t\le T}\|S(t)u_0(\cdot,\omega)\|_{L^1(\mathbb{R}^d)}\big)^k\,\mathbb{P}(d\omega)\\ &= \int_\Omega \big\{\max_{0\le t\le T}\big(\|S(t)u_0(\cdot,\omega)\|^k_{L^1(\mathbb{R}^d)}\big)\big\}\,\mathbb{P}(d\omega)\\ &\le \int_\Omega \|u_0(\cdot,\omega)\|^k_{L^1(\mathbb{R}^d)}\,\mathbb{P}(d\omega) = \|u_0\|^k_{L^k(\Omega;L^1(\mathbb{R}^d))}\,.\end{aligned}$$

This implies (3.20). (3.21) and (3.22) follow from (3.9), (3.10), (3.11).

□

Theorem 3.3 ensures the existence of k-th moments of the random entropy solution $u(x,t;\omega)$ provided that $u_0 \in L^k(\Omega,\mathcal{F},\mathbb{P};L^1(\mathbb{R}^d))$. We next discuss the existence of the (deterministic) k-moments resp. of k-point correlation functions of u.

3.4. **Tensor Products of Banach-spaces.** We have seen in Theorem 3.3 that for initial data $u_0 \in L^k(\Omega,\mathcal{F},\mathbb{P};L^1(\mathbb{R}^d))$ of (3.1) - (3.4), the (unique) random entropy solution $u \in L^k(\Omega;C(0,T;L^1(\mathbb{R}^d)))$ for any $T<\infty$ and the same $k\in\mathbb{N}$. To investigate the existence of the (deterministic) k-th moment (or k-point correlation function) for random entropy solutions, we require facts on tensor products of Banach-spaces; we shall now briefly recapitulate these from [18], Chapter 1, in the form which is best suited for application in the present context.

Let $x,y\in\mathbb{R}^d$, and let $f(x,y)$: $\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ be integrable: $f\in L^1(\mathbb{R}^d\times\mathbb{R}^d;\mathbb{R})$. Then, by Fubini's Theorem, $L^1(\mathbb{R}^d\times\mathbb{R}^d;\mathbb{R})\cong L^1(\mathbb{R}^d;L^1(\mathbb{R}^d;\mathbb{R}))$ and

$$\begin{aligned}\|f\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d,dx\otimes dy)} &= \int_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}|f(x,y)|dxdy\\ = \int_{x\in\mathbb{R}^d}\|f(x,\cdot)\|_{L^1(\mathbb{R}^d,dy)}dx &= \int_{y\in\mathbb{R}^d}\|f(\cdot,y)\|_{L^1(\mathbb{R}^d,dx)}dy\\ = \big\|\,\|f(x,y)\|_{L^1(\mathbb{R}^d,dy)}\big\|_{L^1(\mathbb{R}^d,dx)} &= \big\|\,\|f(x,y)\|_{L^1(\mathbb{R}^d,dx)}\big\|_{L^1(\mathbb{R}^d,dy)}\,.\end{aligned}$$

For $S,T\subseteq\mathbb{R}^d$, we identify for $1\le p<\infty$ the Bochner space $L^p(S,ds;\,L^p(T,dt))$ with $L^p(S\times T,ds\otimes dt)$ and, moreover, we use that for $1\le p<\infty$ these Bochner spaces are isomorphic to tensor product spaces ([18], Chapter 1):

$$L^p(S)\otimes_{\alpha_p} L^p(T)\cong L^p(S\times T),\quad 1\le p<\infty\,. \tag{3.25}$$

Here, α_p indicates the so-called *p-nuclear norm* on the tensor product $X\otimes Y$ of Banach-spaces (see [18], Def. 1.45). In (3.25), we specialize to $X=L^p(S)$, $Y=L^p(T)$ and use [18], Cor. 1.52. In (3.25), $\otimes_{\alpha_p}$ denotes the tensor product space equipped with the p-nuclear norm. We use in the following (3.25) with $p=1$ and denote the tensor product $\otimes_{\alpha_1}$ by $\otimes$. We also remark that (3.25) becomes false for $p=\infty$, in general. The p-nuclear norm (3.25) is a *cross norm*: the norm of simple tensor products ("dyads") $x\otimes y$, $x\in L^p(S)$, $y\in L^p(T)$ satisfies for $1\le p<\infty$

$$\|x\otimes y\|_{L^p(S)\otimes_{\alpha_p}L^p(T)} = \|x\|_{L^p(S)}\|y\|_{L^p(T)}\quad\forall x\in L^p(S),\,y\in L^p(T)\,. \tag{3.26}$$

For $k \in \mathbb{N}$ and separable Banach space X, we denote by $X^{(k)} = \underbrace{X \otimes \cdots \otimes X}_{k \text{ times}}$ the k-fold tensor product of k copies of X. Throughout the following, we shall assume the k-fold tensor product of the Banach-space X with itself, i.e. $X^{(k)}$, to be equipped with a cross norm $\| \circ \|_{X^{(k)}}$ which satisfies, analogous to (3.26),

$$\|u_1 \otimes \cdots \otimes u_k\|_{X^{(k)}} = \|u_1\|_X \ldots \|u_k\|_X \,. \tag{3.27}$$

In particular, for $X = L^p(\mathbb{R}^d)$, $1 \leq p < \infty$, we get the isomorphism

$$L^p(\mathbb{R}^d)^{(k)} \cong L^p(\mathbb{R}^{kd}) \,. \tag{3.28}$$

3.5. **k-th moments.** For $u \in L^k(\Omega; X)$, we consider the random field $(u)^{(k)}$ defined by $\underbrace{u(\omega) \otimes \cdots \otimes u(\omega)}_{k-\text{times}}$. Then

$$(u)^{(k)} = u \otimes \cdots \otimes u \in L^1(\Omega; X^{(k)}) \tag{3.29}$$

and, by (3.27), we have

$$\|(u)^{(k)}\|_{L^1(\Omega; X^{(k)})} = \int_\Omega \|u(\cdot, \omega)\|_X^k \, \mathbb{P}(d\omega) = \|u\|^k_{L^k(\Omega, X)} < \infty \,. \tag{3.30}$$

Therefore, $(u)^{(k)} \in L^1(\Omega, X^{(k)})$ and *the k-the moment* (or k-point correlation function of u)

$$\mathcal{M}^k u := \mathbb{E}[(u)^{(k)}] \in X^{(k)} \tag{3.31}$$

is well-defined for $u \in L^k(\Omega; X)$. With those definitions, we obtain from Theorem 3.3 the following result.

Theorem 3.4. *Consider the scalar conservation law* (3.1) - (3.4) *with random initial data $u_0 : \Omega \to L^1(\mathbb{R}^d)$ satisfying* (3.14), (3.15). *Assume further that for some $k \in \mathbb{N}$, and some real number r such that $k \leq r < \infty$ holds*

$$u_0 \in L^r(\Omega; L^1(\mathbb{R}^d)) \,. \tag{3.32}$$

Then, for every $0 < T < \infty$ and every

$$0 < t_1, t_2, \ldots, t_k \leq T < \infty \tag{3.33}$$

the spatial k-point correlation function

$$u(x_1, t_1; \omega) \otimes \cdots \otimes u(x_k, t_k; \omega) \tag{3.34}$$

is well-defined as an element of $L^{r/k}(\Omega; L^1(\mathbb{R}^{kd}))$. Moreover, the k-th moment

$$(\mathcal{M}^k u)(t_1, \ldots, t_k) := \mathbb{E}[u(\cdot, t_1; \omega) \otimes \cdots \otimes u(\cdot, t_k; \omega)] \tag{3.35}$$

is well-defined for any choice of t_j as in (3.33) *as an element of $L^1(\mathbb{R}^{kd})$, and*

(3.36)

$$\left\| \left(\mathcal{M}^k u\right)(t_1, ..., t_k) \right\|_{L^1(\mathbb{R}^d)^{(k)}} \leq \left\| \bigotimes_{j=1}^k u(\cdot, t_j; \cdot) \right\|_{L^{r/k}(\Omega; (L^1(\mathbb{R}^d))^{(k)})} \leq \|u_0\|^r_{L^r(\Omega; L^1(\mathbb{R}^d))} \,.$$

Proof. By Theorem 3.3, we find from (3.32) that

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

Since $1 \leq k \leq r < \infty$, we find from (3.29), that for $\mathbb{P}$-a.e. $\omega \in \Omega$ and for every $0 < t_1, \ldots, t_k < T$ holds

$$u(\cdot, t_1; \omega) \otimes \cdots \otimes u(\cdot, t_k; \omega) \in L^{r/k}(\Omega; L^1(\mathbb{R}^d)^{(k)}) \cong L^{r/k}(\Omega; L^1(\mathbb{R}^{kd})) \tag{3.37}$$

which is (3.34). Since $r/k \geq 1$, the k-th moment

$$(\mathcal{M}^k u)(t_1, \ldots, t_k) = \int_\Omega \bigotimes_{j=1}^k u(\cdot, t_j; \omega)\, \mathbb{P}(d\omega) \tag{3.38}$$

satisfies, using the cross-norm property (3.27), the bounds

$$\begin{aligned}
\|(\mathcal{M}^k u)(t_1, \ldots, t_k)\|_{L^1(\mathbb{R}^{kd})} &= \Big\| \int_\Omega \bigotimes_{j=1}^k u(\cdot, t_j; \omega)\, \mathbb{P}(d\omega) \Big\|_{L^1(\mathbb{R}^{kd})} \\
&\leq \int_\Omega \Big\| \bigotimes_{j=1}^k u(\cdot, t_j; \omega) \Big\|_{L^1(\mathbb{R}^{kd})} \mathbb{P}(d\omega) \\
&= \int_\Omega \prod_{j=1}^k \|u(\cdot, t_j; \omega)\|_{L^1(\mathbb{R}^{kd})}\, \mathbb{P}(d\omega) \\
&\leq \int_\Omega \|u_0(\cdot; \omega)\|^k_{L^1(\mathbb{R}^d)}\, \mathbb{P}(d\omega) \\
&= \|u_0\|_{L^k(\Omega;L^1(\mathbb{R}^d))} < \infty
\end{aligned}$$

by (3.32). Hence (3.38) is well-defined as element in $(L^1(\mathbb{R}^d))^{(k)} \cong L^1(\mathbb{R}^{kd})$, and (3.36), (3.37) follow from Hölder's inequality since $r/k \geq 1$ by assumption. □

4. Multilevel Monte Carlo Finite Volume Method

4.1. **Monte-Carlo Method.** We view the Monte-Carlo Method as a "discretization" of the hyperbolic IVP (3.1) - (3.4) with random initial data $u_0(x; \omega)$ as in (3.13) - (3.15) with respect to ω. We also assume (3.16), i.e. the existence of k-th moments of u_0 for some $k \in \mathbb{N}$, to be specified later. We shall be interested in the statistical estimation of the first and higher moments of u i.e, $\mathcal{M}^k(u) \in (L^1(\mathbb{R}^d))^{(k)}$. For $k = 1$, $\mathcal{M}^1(u) = \mathbb{E}[u]$. The *MC approximation of* $\mathbb{E}[u]$ is defined as follows: given M independent, identically distributed samples $\widehat{u}_0^i$, $i = 1, \ldots, M$, of initial data, the MC estimate of $\mathbb{E}[u(\cdot, t; \cdot)]$ at time t is given by

$$E_M[u(\cdot, t)] := \frac{1}{M} \sum_{i=1}^M \widehat{u}^i(\cdot, t) \tag{4.1}$$

where $\widehat{u}^i(\cdot, t)$ denotes the M unique entropy solutions of the M Cauchy Problems (3.1) - (3.4) with initial data $\widehat{u}_0^i$. We observe that by

$$\widehat{u}^i(\cdot, t) = S(t)\, \widehat{u}_0^i \tag{4.2}$$

we have from (3.9) - (3.11) for every M and for every $0 < t < \infty$, by (3.11),

(4.3)
$$\|E_M[u(\cdot,t;\omega)]\|_{L^1(\mathbb{R}^d)} = \Big\| \frac{1}{M} \sum_{i=1}^{M} S(t)\, \widehat{u}_0^i(\cdot;\omega) \Big\|_{L^1(\mathbb{R}^d)} \le \frac{1}{M} \sum_{i=1}^{M} \big\| S(t)\, \widehat{u}_0^i(\cdot;\omega) \big\|_{L^1(\mathbb{R}^d)}$$
$$\le \frac{1}{M} \sum_{i=1}^{M} \|\widehat{u}_0^i(\cdot;\omega)\|_{L^1(\mathbb{R}^d)} .$$

Using the i.i.d. property of the samples $\{\widehat{u}_0^i\}_{i=1}^M$ of the random initial data u_0, Lemma 2.1 and the linearity of the expectation $\mathbb{E}[\cdot]$, we obtain the bound

$$\mathbb{E}\big[\|E_M[u(\cdot,t)]\|_{L^1(\mathbb{R}^d)}\big] \le \mathbb{E}\big[\|u_0\|_{L^1(\mathbb{R}^d)}\big] = \|u_0\|_{L^1(\Omega;L^1(\mathbb{R}^d))} < \infty . \tag{4.4}$$

As $M \to \infty$, the MC estimates (4.1) converge and we obtain the following convergence result.

Theorem 4.1. *Assume that in* (3.1) - (3.4) *the random initial data* u_0 *satisfies*

$$u_0 \in L^2(\Omega; L^1(\mathbb{R}^d)) \tag{4.5}$$

and that (3.14), (3.15) *hold. Then the MC estimates* $E_M[u(\cdot,t)]$ *in* (4.1) *converge as* $M \to \infty$, *to* $\mathcal{M}^1(u(\cdot,t)) = \mathbb{E}[u(\cdot,t)]$ *and, for any* $M \in \mathbb{N}$, $0 < t < \infty$, *there holds the error bound*

$$\|\mathbb{E}[u(\cdot,t)] - E_M[u(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \le 2M^{-\frac{1}{2}} \|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} . \tag{4.6}$$

Proof. As is customary in the convergence analysis of MC methods, we interpret the M samples $\{\widehat{u}_0^i\}_{i=1}^M$ as realizations of M independent "copies" of u_0 on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $\{\widehat{u}_0^i\}_{i=1}^M$ are M i.i.d. copies of $u_0 \in L^2(\Omega; L^1(\mathbb{R}^d))$. The corresponding unique random entropy solutions $\widehat{u}^i(\cdot,t;\omega) = S(t)\,\widehat{u}_0^i(\cdot;\omega)$, $i = 1,\dots,M$ are then also independent in $L^2(\Omega; C(0,T; L^1(\mathbb{R}^d)))$ by (3.21):

the images of any two i.i.d. realizations of $u_0(x;\omega)$ under the (deterministic, nonlinear) solution map $S(t)$ are, for any fixed $t > 0$, strongly measurable as $L^1(\mathbb{R}^d)$-valued functions by the $L^1(\mathbb{R}^d)$ contractivity (3.8). By Lemma 2.1 and by the continuity (3.11), for every $0 < t < \infty$, the mapping $\Omega \ni \omega \to \|u(\cdot,t;\omega)\|_{L^1(\mathbb{R}^d)}$ is $\mathbb{P}$-measurable. Since

$$\sup_{0\le t\le T} \int_\Omega \|u(\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\, \mathbb{P}(d\omega) \le \int_\Omega \sup_{0\le t\le T} \|u(\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\, \mathbb{P}(d\omega)$$
$$= \|u\|^2_{L^2(\Omega;C(0,T;L^1(\mathbb{R}^d)))} < \infty , \tag{4.7}$$

also $u \in C([0,T];\, L^2(\Omega; L^1(\mathbb{R}^d)))$. Next, we calculate for $0 < t < \infty$ and any $M \in \mathbb{N}$ with (4.2) and (3.19) and with the notation $\bar{u}(\cdot,t) = \mathcal{M}^1[u(\cdot,t)] = \mathbb{E}[u(\cdot,t;\omega)]$,

$$\mathbb{E}\big[\|\bar{u}(\cdot,t) - E_M[u]\|^2_{L^1(\mathbb{R}^d)}\big] = \mathbb{E}\Big[\frac{1}{M^2}\Big\| \sum_{i=1}^{M} (\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\Big\|^2_{L^1(\mathbb{R}^d)}\Big]$$

leading to

$$\mathbb{E}\big[\|(\bar{u}(\cdot,t) - E_M[u](\cdot,t;\omega))\|^2_{L^1(\mathbb{R}^d)}\big] \le \mathbb{E}\Big[M^{-2}\Big(\sum_{i=1}^{M} \|(\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\|_{L^1(\mathbb{R}^d)}\Big)^2\Big] . \tag{4.8}$$

Since $L^1(\mathbb{R}^d)$ is separable, and since $(L^1(\mathbb{R}^d))^* \cong L^\infty(\mathbb{R}^d)$, for $i = 1, \dots, M$, every $0 < t < \infty$ and for every $\omega \in \Omega$ exist unique $\varphi_i(\cdot, t; \omega) \in L^\infty(\mathbb{R}^d)$, $i = 1, ..., M$, such that

(4.9)

$$\|\varphi_i(\cdot,t;\omega)\|_{L^\infty(\mathbb{R}^d)} = 1, \quad \|(\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\|_{L^1(\mathbb{R}^d)} = \big\langle \varphi_i(\cdot,t;\omega), (\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\big\rangle$$

where $\langle \cdot, \cdot \rangle$ denotes the (L^∞, L^1) duality pairing. We expand the square in (4.8) to get

$$\mathbb{E}\big[\|\bar{u}(\cdot,t) - E_M[u](\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\big] \le \frac{1}{M^2}\mathbb{E}\Big[\Big(\sum_{i=1}^M \big\langle \varphi_i(\cdot,t;\omega), (\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\big\rangle\Big)^2\Big]$$

$$= \frac{1}{M^2}\mathbb{E}\Big[\sum_{i,j=1}^M \big\langle \varphi_i(\cdot,t;\omega), (\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\big\rangle \big\langle \varphi_j(\cdot,t;\omega), (\bar{u}(\cdot,t) - \widehat{u}^j(\cdot,t;\omega))\big\rangle\Big].$$

Using the linearity of the expectation $\mathbb{E}[\cdot]$, and the fact that independence of the samples $\widehat{u}^i_0(\cdot,\omega) \in L^1(\mathbb{R}^d)$ implies independence of the associated random entropy solutions $\widehat{u}^i(\cdot,t;\omega) = S(t)\widehat{u}^i_0(\cdot,\omega)$, we infer that also the representers $\varphi_i(\cdot,t;\omega) \in L^\infty(\mathbb{R}^d)$ are independent. We get with (4.9) that M^2 times the last expression equals

$$\mathbb{E}\Big[\sum_{i=1}^M \big|\big\langle \varphi_i(\cdot,t;\omega), (\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega))\big\rangle\big|^2\Big] = \mathbb{E}\Big[\sum_{i=1}^M \|\bar{u}(\cdot,t) - \widehat{u}^i(\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\Big].$$

Using that $u_0 \in L^2(\Omega; L^1(\mathbb{R}^d))$ and that the $\widehat{u}^i_0$ are identically distributed to u_0, we obtain with the elementary inequality $\|x - y\|^2 \le 2(\|x\|^2 + \|y\|^2)$ and the apriori estimate (3.21) in Theorem 3.3 the bound

$$\mathbb{E}\big[\|\bar{u}(\cdot,t) - E_M[u](\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\big] \le \frac{2}{M^2}\Big(\sum_{i=1}^M \big(\|\bar{u}(\cdot,t)\|^2_{L^1(\mathbb{R}^d)} + \mathbb{E}\big[\|\widehat{u}^i(\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\big]\big)\Big).$$

Using here the apriori estimates (3.36) with $k = r = 1$ and (3.11) gives

$$\begin{aligned} \|\mathbb{E}[u(\cdot,t;\omega)]\|_{L^1(\mathbb{R}^d)} &\le \mathbb{E}\Big[\|u(\cdot,t;\omega)\|_{L^1(\mathbb{R}^d)}\Big] = \mathbb{E}\Big[\|S(t)u_0(\cdot;\omega)\|_{L^1(\mathbb{R}^d)}\Big] \\ &\le \mathbb{E}\Big[\|u_0(\cdot;\omega)\|_{L^1(\mathbb{R}^d)}\Big] \end{aligned}$$

and we arrive at

$$\mathbb{E}\big[\|\bar{u}(\cdot,t) - E_M[u](\cdot,t;\omega)\|^2_{L^1(\mathbb{R}^d)}\big] \le 4M^{-1}\,\mathbb{E}\big[\|u_0(\cdot,\omega)\|^2_{L^1(\mathbb{R}^d)}\big]$$

which implies (4.6) upon taking square roots. □

So far, we addressed the MC estimation of the mean field or first moment. A similar result holds for the MC estimates of the k-th moment $\mathcal{M}^k u := \mathbb{E}[(u)^{(k)}] \in (L^1(\mathbb{R}^d))^{(k)}$.

Theorem 4.2. *Consider the scalar conservation law* (3.1) - (3.4) *with random initial data* $u_0 : \Omega \to L^1(\mathbb{R}^d)$ *satisfying* (3.14), (3.15). *Assume further that for some* $k \in \mathbb{N}$ *holds* $u_0 \in L^{2k}(\Omega; L^1(\mathbb{R}^d))$. *Then the MC estimates*

$$E_M[(u(\cdot,t)^{(k)}] := \frac{1}{M}\sum_{i=1}^M (\widehat{u}^i(\cdot,t))^{(k)} \tag{4.10}$$

with the M i.i.d samples $\widehat{u}^i(\cdot,t)$, $i=1,2,...$, converge to the k-th moment (or spatial k-point correlation function) $(\mathcal{M}^k u)(t)$ defined in (3.38)*. Moreover, there hold the error bounds*

$$\|(\mathcal{M}^k u)(t) - E_M[(u(\cdot,t;\omega))^{(k)}]\|_{L^2(\Omega;L^1(\mathbb{R}^{kd}))} \leq 2M^{-1/2}\|u_0\|^k_{L^{2k}(\Omega;L^1(\mathbb{R}^d))}\,. \tag{4.11}$$

Proof. The assumption that $u_0 \in L^{2k}(\Omega;L^1(\mathbb{R}^d))$ implies in (3.36) of Theorem 3.4 (with $r=2k$) and (3.28) with $p=1$ that, for any choice of time instants $0<t_1,...,t_k<\infty$ and for $\mathbb{P}$-a.e. $\omega\in\Omega$, the spatial k point correlation

$$\bigotimes_{j=1}^{k} u(\cdot,t_j;\omega) \in (L^1(\mathbb{R}^d))^k$$

of the random entropy solution $u(x,t;\omega)$ is well-defined in $L^2(\Omega;L^1(\mathbb{R}^{kd}))$. Hence the proof of Theorem 3.4 directly applies to $(u(\cdot,t;\omega))^{(k)}$. □

4.2. **Finite Volume Method (FVM).** So far, we considered the MCM under the assumption that the entropy solutions $\widehat{u}^i(x,t;\omega)=S(t)\,\widehat{u}^i_0(x;\omega)$ for the Cauchy problem (3.1) - (3.4) with the initial data samples $\hat{u}^i_0$ are available exactly. In practice, however, numerical approximations of $S(t)\hat{u}^i_0$ must be computed by FVM. We analyze the error of the combined MC-FVM approximations. In order to simplify the exposition, we consider only first-order FVM in this section.

We assume given a time step $\Delta t>0$ and a triangulation $\mathcal{T}$ of the spatial domain $D\subset\mathbb{R}^d$ of interest. Here, a triangulation $\mathcal{T}$ will be understood as a set of open, convex polyhedra $K\subset\mathbb{R}^d$ with plane faces such that the following conditions hold: the triangulation $\mathcal{T}$ is *shape regular*: if $K\in\mathcal{T}$ denotes a generic volume, we define the volume parameter

$$\rho_K=\rho(K)=\max\{\mathrm{diam}(B_r):\ B_r\subset\overline{K}\} \tag{4.12}$$

i.e., the maximum diameter of balls B_r of radius $r>0$ that can be inscribed into volume $K\in\mathcal{T}$ and define, in addition, for a generic mesh $\mathcal{T}$, the *shape regularity constants* (where $\Delta x_K:=\mathrm{diam}\,K$)

$$\kappa(\mathcal{T}):=\sup\{\Delta x_K/\rho(K):\ K\in\mathcal{T}\},\ \mathcal{T}\in\mathfrak{M}\,. \tag{4.13}$$

We also denote by $\Delta x(\mathcal{T}):=\max\{\Delta x_K:\ K\in\mathcal{T}\}$ the *mesh width* of $\mathcal{T}$. For any volume $K\in\mathcal{T}$, we define the *set $\mathcal{N}(K)$ of neighboring volumes*

$$\mathcal{N}(K):=\{K'\in\mathcal{T}:K'\neq K\wedge \mathrm{meas}_{d-1}(\overline{K}\cap\overline{K'})>0\}\,. \tag{4.14}$$

We assume that the triangulation $\mathcal{T}$ are regular in the sense that there exists an absolute constant $B>0$ independent of $\Delta x(\mathcal{T})$ such that the support size of the FV "stencil" at element $K\in\mathcal{T}$ is uniformly bounded

$$\sigma(\mathcal{T}):=\sup_{K\in\mathcal{T}}\#(\mathcal{N}(K))\leq B\,. \tag{4.15}$$

We set

$$\lambda=\Delta t/\Delta x(\mathcal{T})\,. \tag{4.16}$$

by assuming a uniform discretization in time with time step Δt. The constant lambda is determined by a standard CFL condition (see [9]) based on the maximum wave speed.

To approximate (3.1), we use a time-explicit, first order FV scheme on $\mathcal{T}$. It has the generic form

$$v_K^{n+1} = H(\{v_{K'}^n : K' \in \mathcal{N}(K) \cup K\}), \quad K \in \mathcal{T} \tag{4.17}$$

where $H : \mathbb{R}^{(2k+1)^d} \to \mathbb{R}$ 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 $H_{\mathcal{T}}((v_K)_{K\in\mathcal{T}})$. Then (4.17) takes the abstract form

$$\underline{v}^{n+1} = H_{\mathcal{T}}(\underline{v}^n), \quad n = 0, 1, 2, \dots . \tag{4.18}$$

For the ensuing convergence analysis, we shall assume and use several properties of the FV scheme (4.18); these properties are satisfied by many commonly use FVM of the form (4.18), 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 $(v_K^0)_{K\in\mathcal{T}}$ by the *cell averages*

$$v_K^0 = \frac{1}{|K|} \int_K u_0(x)\, dx, \quad \text{where } K \in \mathcal{T}\,. \tag{4.19}$$

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

$$v_{\mathcal{T}}(x,t)\big|_K := v_K, \quad K \in \mathcal{T}\,. \tag{4.20}$$

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:

$$\|\underline{v}\|_{L^1(\mathcal{T})} = \sum_{K\in\mathcal{T}} |K||v_K| \ = \|v_{\mathcal{T}}\|_{L^1(\mathbb{R}^d)}\,, \tag{4.21}$$

$$\|\underline{v}\|_{L^\infty(\mathcal{T})} = \sup_{K\in\mathcal{T}} |v_K| = \|v_{\mathcal{T}}\|_{L^\infty(\mathbb{R}^d)}\,. \tag{4.22}$$

For any function $v \in L^1(\mathbb{R}^d)$ and any triangulation $\mathcal{T}$, the linear mapping $v \to v_{\mathcal{T}}$ induced by (4.19), (4.20) is denoted by $P_{\mathcal{T}}$:

$$v_{\mathcal{T}} = P_{\mathcal{T}} v, \quad P_{\mathcal{T}} : L^1(\mathbb{R}^d) \to S(\mathcal{T}). \tag{4.23}$$

Then $P_{\mathcal{T}}^2 v = P_{\mathcal{T}} v$ and

$$\forall v \in (L^1 \cap L^\infty)(\mathbb{R}^d): \quad \|P_{\mathcal{T}} v\|_{L^1(\mathbb{R}^d)} \le \|v\|_{L^1(\mathbb{R}^d)}, \quad \|P_{\mathcal{T}} v\|_{L^\infty(\mathbb{R}^d)} \le \|v\|_{L^\infty(\mathbb{R}^d)}\,. \tag{4.24}$$

We also have for every $v \in W^{s,1}(\mathbb{R}^d)$ for some $0 \le s \le 1$ the approximation property

$$\|v - P_{\mathcal{T}} v\|_{L^1(\mathbb{R}^d)} \le C(\Delta x(\mathcal{T}))^s |v|_{W^{s,1}(\mathbb{R}^d)}\,. \tag{4.25}$$

where

$$\Delta x(\mathcal{T}) = \sup\{\mathrm{diam}(K) : \ K \in \mathcal{T}\} \tag{4.26}$$

denotes the mesh width of $\mathcal{T}$. We shall assume the following properties of the FVM schemes used in the MC-FVM algorithms.

Assumption 4.3. *We shall assume that the abstract FV scheme* (4.18) *satisfies*

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

$$\|v_{\mathcal{T}}(\cdot,t)\|_{L^\infty(\mathbb{R}^d)} \leq \|v_{\mathcal{T}}(\cdot,0)\|_{L^\infty(\mathbb{R}^d)}, \tag{4.27}$$

$$\|v_{\mathcal{T}}(\cdot,t)\|_{L^1(\mathbb{R}^d)} \leq \|v_{\mathcal{T}}(\cdot,0)\|_{L^1(\mathbb{R}^d)}, \tag{4.28}$$

$$TV(v_{\mathcal{T}}(\cdot,t)) \leq TV(v_{\mathcal{T}}(\cdot,0)), \tag{4.29}$$

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*

$$\|H_{\mathcal{T}}(\underline{v}) - H_{\mathcal{T}}(\underline{w})\|_{L^1(\mathcal{T})} \leq \|\underline{v} - \underline{w}\|_{L^1(\mathcal{T})} \tag{4.30}$$

or, equivalently,

$$\|H_{\mathcal{T}}(v_{\mathcal{T}}) - H_{\mathcal{T}}(w_{\mathcal{T}})\|_{L^1(\mathbb{R}^d)} \leq \|v_{\mathcal{T}} - w_{\mathcal{T}}\|_{L^1(\mathbb{R}^d)}\,. \tag{4.31}$$

3. Convergence: *If* $\lambda = \Delta t/\Delta x$ *is kept constant, as* $\Delta x \to 0$, *the approximate solution* $v_\Delta(x,t)$ *generated by* (4.17) - (4.20) *converges to the unique entropy solution* u *of the scalar conservation laws* (3.1) - (3.4) *at rate* $0 < s \leq 1$, *i.e., there exists* $C > 0$ *independent of* Δx *such that, as* $\Delta x \to 0$, *for every* $\bar{t}$ *such that, for* $(\Delta t)^s \leq \bar{t} \leq T$, *it holds*

$$\|u(\cdot,\bar{t}) - v_{\mathcal{T}}(\cdot,\bar{t})\|_{L^1(\mathbb{R}^d)} \leq \|u_0 - v^0_{\mathcal{T}}\|_{L^1(\mathbb{R}^d)} + C\,\bar{t}\,TV(u_0)\,\Delta t^s\,. \tag{4.32}$$

Let us mention that Assumption 4.3 is satisfied for many standard FVM-schemes; we refer to [9, 10, 17] and the references there for further details. Let us also mention that the work for the realization of scheme (4.17) - (4.20) on a bounded domain $D \subset \mathbb{R}^d$ as (using the CFL stability condition (4.16), i.e. $\Delta t/\Delta x \leq \lambda = \text{const.}$)

$$\text{Work}_{\mathcal{T}} = O(\Delta t^{-1}\,\Delta x^{-d}) = O(\Delta x^{-(d+1)})\,. \tag{4.33}$$

The convergence estimate (4.32) is known to hold for first-order FVM by results of Kusznetsov (see, e.g. [11]) with $s = 1/2$.

In the Monte Carlo Finite Volume Methods (MC-FVMs), we combine MC sampling of the random initial data with the FVM (4.18). In the convergence analysis of these schemes, we shall require the application of the FVM (4.18) to random initial data $u_0 \in L^\infty(\Omega; L^1(\mathbb{R}^d))$. Given a draw $u_0(x;\omega)$ of u_0, the FVM (4.18) - (4.20) defines a family $v_{\mathcal{T}}(x,t;\omega)$ of grid functions. There holds

Proposition 4.4. *Consider the FVM* (4.18) - (4.20) *for the approximation of the entropy solution corresponding to the draw* $u_0(x;\omega)$ *of the random initial data.*

Then, if the FVM satisfies Assumption 4.3, the random grid functions $\Omega \ni \omega \longmapsto v_{\mathcal{T}}(x,t;\omega)$ *defined by* (4.16) - (4.20) *satisfy, for every* $0 < \bar{t} < \infty$, $0 < \Delta x < 1$, *and every* $k \in \mathbb{N} \cup \{\infty\}$ *the stability bounds:*

$$\|v_{\mathcal{T}}(\cdot,\bar{t};\cdot)\|_{L^k(\Omega;L^\infty(\mathbb{R}^d))} \leq \|u_0\|_{L^k(\Omega;L^\infty(\mathbb{R}^d))}\,, \tag{4.34}$$

$$\|v_{\mathcal{T}}(\cdot,\bar{t};\cdot)\|_{L^k(\Omega;L^1(\mathbb{R}^d))} \leq \|u_0\|_{L^k(\Omega;L^1(\mathbb{R}^d))}\,. \tag{4.35}$$

There also holds the error bound

$$\begin{aligned} &\|u(\cdot,\bar{t};\omega) - v_{\mathcal{T}}(\cdot,\bar{t};\omega)\|_{L^k(\Omega;L^1(\mathbb{R}^d))} \\ \leq\ &\|u_0(\cdot;\omega) - v^0_{\mathcal{T}}(\cdot;\omega)\|_{L^k(\Omega;L^1(\mathbb{R}^d))} + C\bar{t}\Delta t^s\,\|TV(u_0(\cdot;\omega)\|_{L^k(\Omega,d\mathbb{P})}\,. \end{aligned} \tag{4.36}$$

4.3. **MC-FVM Scheme.** We next define and analyze the MC-FVM scheme. It is based on the straightforward idea of generating, possibly in parallel, independent samples of the random initial data and then, for each sample of the random initial data, to perform one FV simulation. The error of this procedure is bound by two contributions: a (statistical) sampling error and a (deterministic) discretization error. We express the asymptotic efficiency of this approach (in terms of overall error versus work). It will be seen that the efficiency of the MC-FVM is, in general, inferior to that of the deterministic scheme (4.18). The present analysis will constitute a key technical tool in our subsequent development and analysis of the multilevel MC-FVM ("MLMC-FVM" for short) which does not suffer from this drawback.

4.3.1. *Definition of the MC-FVM Scheme.* We consider once more the initial value problem (3.1) - (3.4) with random initial data u_0 satisfying (3.13) - (3.16) for sufficiently large $k \in \mathbb{N}$ (to be specified in the convergence analysis). The MC-FVM scheme for the MC estimation of the mean of the random entropy solutions then consists in the following:

Definition 4.5. *(MC-FVM Scheme)*

Given $M \in \mathbb{N}$, generate M i.i.d. samples $\{\widehat{u}_0^i\}_{i=1}^M$ of initial data. Let $\{\widehat{u}^i(\cdot,t)\}_{i=1}^M$ denote the unique entropy solutions of the scalar conservation laws (3.1) - (3.4) *for these data samples, i.e.*

$$\widehat{u}^i(\cdot,t) = S(t)\,\widehat{u}_0^i(\cdot), \quad i = 1,\dots,M\,. \tag{4.37}$$

Let $H_{\mathcal{T}}(\cdot)$ be a FVM scheme (4.17) - (4.20) *satisfying Assumption* 4.3. *Then the MC-FVM approximations of $\mathcal{M}^k(u(\cdot,t))$ are defined as statistical estimates from the ensemble*

$$\{\widehat{v}_{\mathcal{T}}^i(\cdot,t)\}_{i=1}^M \tag{4.38}$$

obtained by (4.18) *from the FV approximations $\widehat{v}_{\mathcal{T}}^i(\cdot,0)$ of the initial data $\{\widehat{u}_0^i(x)\}_{i=1}^M$ samples by* (4.19)*: specifically, the first moment of the random solution $u(\cdot,t;\omega)$ at time $t>0$, is estimated as*

$$\mathcal{M}^1(u(\cdot,t)) \approx E_M[v_{\mathcal{T}}(\cdot,t)] := \frac{1}{M}\sum_{i=1}^M \widehat{v}_{\mathcal{T}}^i(\cdot,t)\,, \tag{4.39}$$

and, for $k>1$, the kth moment (or k point correlation function) $\mathcal{M}^k(u(\cdot,t)) = \mathbb{E}[(u(\cdot,t))^{(k)}]$ defined in (3.35) *is estimated by*

$$E_M^{(k)}[v_{\mathcal{T}}(\cdot,t)] := \frac{1}{M}\sum_{i=1}^M \underbrace{(\widehat{v}_{\mathcal{T}}^i \otimes \cdots \otimes \widehat{v}_{\mathcal{T}}^i)}_{k-\text{times}}(\cdot,t)\,. \tag{4.40}$$

More generally, for $k>1$, we consider time instances $t_1,\dots,t_k \in (0,T]$, $T<\infty$, and define the statistical FVM estimate of $\mathcal{M}^k(u)(t_1,...,t_k)$ by

$$E_M^{(k)}\left[v_{\mathcal{T}}\right](t_1,\dots,t_k) := \frac{1}{M}\sum_{i=1}^M \underbrace{(\widehat{v}_{\mathcal{T}}^i(\cdot,t_1) \otimes \cdots \otimes \widehat{v}_{\mathcal{T}}^i(\cdot,t_k))}_{k-\text{times}}\,. \tag{4.41}$$

4.3.2. *Convergence Analysis of MC-FVM.* We next address the convergence of $E_M[v_{\mathcal{T}}]$ to the mean $\mathbb{E}(u)$.

Theorem 4.6. *Assume that*

$$u_0 \in L^\infty(\Omega, L^1(\mathbb{R}^d)) \tag{4.42}$$

and that (3.13) - (3.15) *hold. Assume further that we are given a FVM* (4.17) - (4.20) *such that* (4.16) *holds and such that Assumption* 4.3 *is satisfied; in particular, assume that the deterministic FVM scheme converges at rate $s > 0$ in $L^\infty([0,T]; L^1(\mathbb{R}^d))$ for every $0 < T < \infty$. Then the MC estimate $E_M[v_{\mathcal{T}}(\cdot,t)]$ defined in* (4.39) *satisfies, for every M, the error bound*

(4.43)

$$\|\mathbb{E}[u(\cdot,t)] - E_M[v_{\mathcal{T}}(\cdot,t;\omega)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \leq$$
$$C\big\{M^{-\frac{1}{2}}\|u_0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} + \|u_0 - v^0_{\mathcal{T}}\|_{L^\infty(\Omega;L^1(\mathbb{R}^d))} + t\Delta t^s \left\|TV(u_0(\cdot,\omega))\right\|_{L^\infty(\Omega;d\mathbb{P})}\big\}$$

where $C > 0$ is independent of M and of Δt as $M \to \infty$ and as $\lambda\Delta x = \Delta t \downarrow 0$. The convergence rate $\Delta x^s > 0$ is as in (4.32).

Proof. We estimate, for arbitrary $t > 0$,

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

Term I is bounded by (4.6). For Term II, we note that, by (4.42) and by (3.9) - (3.12) and Assumption 4.3 with the notation $\bar{u}(\cdot,t) = \mathbb{E}[u(\cdot,t;\omega)]$, by the triangle inequality that

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

□

4.3.3. *Work estimates.* For computational purposes, we have to assume that the computational domain $D \subset \mathbb{R}^d$ is bounded and suitable boundary conditions are specified on ∂D. Noting that in a bounded domain D, the work for one time step (4.17), (4.18) is of order $O(\Delta x^{-d})$, (with $O(\cdot)$ depending on the size of the domain), we find from the CFL condition (4.16) that the total computational work to obtain $\{v_{\mathcal{T}}(\cdot,t)\}_{0<t\leq T}$ in D is by (4.33)

$$\mathrm{Work}(\mathcal{T}) = O(\Delta x^{-d-1}),\ \lambda\Delta x = \Delta t \downarrow 0 \tag{4.44}$$

which implies that the work for the computation of the MC estimate $E_M[v_{\mathcal{T}}(\cdot,t)]$ is

$$\mathrm{Work}(M,\mathcal{T}) = O(M\Delta x^{-d-1}),\ \text{ as }\ \Delta t = \lambda\Delta x \downarrow 0\,, \tag{4.45}$$

so that we obtain from (4.43) the convergence order in terms of work: to this end we equilibrate in (4.43) the two bounds by choosing $M^{-1/2} \sim \Delta t^s$, i.e. $M = \Delta t^{-2s}$. Inserting in (4.45) yields

$$\text{Work}(\mathcal{T}) = O(\Delta t^{-2s}\, \Delta x^{-(d+1)}) \overset{(4.16)}{=} O(\Delta x^{-(d+1)-2s}) \tag{4.46}$$

so that we obtain from (4.43)

$$\|\mathbb{E}[u(\cdot,t)] - E_M[v_{\mathcal{T}}(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \leq C\Delta t^s \leq C(\text{Work}(\mathcal{T}))^{-s/(d+1+2s)}\,. \tag{4.47}$$

We sum up the foregoing considerations.

Remark 4.7. (Work vs. accuracy of MC-FVM)

Let us add some comments on the exponent in (4.47). In the *deterministic FV scheme*, we obtain

$$\text{Work}(\mathcal{T}) = O(\Delta t^{-1}\, \Delta x^{-d}) \overset{(4.16)}{=} O(\Delta x^{-(d+1)}),$$

and the error in terms of work bound (4.32) becomes

$$\|u(\cdot,\bar{t}) - v_{\mathcal{T}}(\cdot,\bar{t})\|_{L^1(\mathbb{R}^d)} \leq \|u_0 - v^0_{\mathcal{T}}\|_{L^1(\mathbb{R}^d)} + C\bar{t}\, TV(u_0)\ (\text{Work}(\mathcal{T}))^{-s/(d+1)}\,. \tag{4.48}$$

Assuming exact representation of the initial data, we obtain the exponent $-s/(d+1)$ for the deterministic FVM as compared to $-s/(d+1+2s)$ for the MC-FVM. We see in particular in the (typical) situation of low order s of convergence and space dimension $d = 2, 3$ a considerably reduced rate of convergence of the MC-FVM, in terms of accuracy vs. work, is obtained. On the other hand, for high order schemes (i.e. when $s >> d+1$) the MC error dominates and we recover in (4.48) the rate $1/2$ in terms of work which is typical of MC methods.

4.4. **Multilevel MC-FVM.** We next present and analyze a scheme that allows us to achieve almost the accuracy versus work bound (4.48) of the deterministic FVM also for the stochastic initial data u_0, rather than the single level MC-FVM error bound (4.47). The key ingredient in the Multilevel Monte Carlo Finite Volume (MLMC-FVM) scheme is simultaneous MC sampling on different levels of resolution of the FVM, *with level dependent numbers M_ℓ of MC samples*. To define these, we introduce some notation.

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

$$\Delta x_\ell = \Delta x(\mathcal{T}_\ell) = \sup\{\text{diam}(K) :\ K \in \mathcal{T}_\ell\} = O(2^{-\ell}\Delta x_0),\ \ \ell \in \mathbb{N}_0 \tag{4.49}$$

where K denotes a generic finite volume 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 (4.13). 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 insribed into K

$$\kappa(\mathfrak{M}) = \sup_{\mathcal{T}\in\mathfrak{M}} \kappa(\mathcal{T}) = \sup_{\mathcal{T}\in\mathfrak{M}} \sup_{K\in\mathcal{T}} \frac{\text{diam}(K)}{\rho_K}\,. \tag{4.50}$$

We recall from (4.23) the definition of the cell-average projections $P_{\mathcal{T}}$ onto $S(\mathcal{T})$. For a mesh hierarchy $\mathfrak{M} = \{\mathcal{T}_\ell\}_{\ell=0}^\infty$, we denote

$$S_\ell := S(\mathcal{T}_\ell), \quad P_\ell := P_{\mathcal{T}_\ell}, \qquad \mathcal{T}_\ell \in \mathfrak{M}, \quad \ell = 0, 1, \dots . \tag{4.51}$$

4.4.2. *Derivation of MLMC-FVM.* As in plain MC-FVM, our aim is to estimate, for $0 < t < \infty$, the expectation $\mathbb{E}[u(\cdot,t)]$ of the random entropy solution of the SCL (3.1) - (3.4) with random initial data $u_0(\cdot,\omega)$, $\omega \in \Omega$, satisfying (3.13) - (3.16) for sufficiently large values of k (to be specified in the sequel). As in the previous section, $\mathbb{E}[u(\cdot,t)]$ will be estimated by replacing $u(\cdot,t)$ by a FVM approximation. For $\ell \in \mathbb{N}_0$, we denote in the present section the FV approximation $v_{\mathcal{T}}$ by $v_\ell(\cdot,t)$ on mesh $\mathcal{T}_\ell \in \mathfrak{M}$, where we assume that the CFL condition (4.16) takes the form

$$\Delta t_\ell \le \lambda \Delta x_\ell, \quad \ell = 0, 1, 2, \dots, \tag{4.52}$$

with a constant $\lambda > 0$ which is independent of ℓ.

By the stability of the FVM scheme, we generate a sequence of stable approximations, $\{v_\ell(\cdot,t)\}_{\ell=0}^\infty$ on $\mathcal{T}_\ell$ for a number of time steps of sizes Δt_ℓ adapted to grid $\mathcal{T}_\ell \in \mathfrak{M}$. We set in what follows $v_{-1}(\cdot,t) := 0$. Then, given a level $L \in \mathbb{N}$ of spatial resolution, we may write by the linearity of the expectation operator

$$\mathbb{E}[v_L(\cdot,t)] = \mathbb{E}\Big[\sum_{\ell=0}^{L} (v_\ell(\cdot,t) - v_{\ell-1}(\cdot,t))\Big] . \tag{4.53}$$

We next estimate each term in (4.53) statistically by a MCM with a level-dependent number of samples, M_ℓ; this gives the MLMC estimator

$$E^L[u(\cdot,t)] = \sum_{\ell=0}^{L} E_{M_\ell}[v_\ell(\cdot,t) - v_{\ell-1}(\cdot,t)] \tag{4.54}$$

where $E_M[v_\Delta(\cdot,t)]$ is as in (4.39), and where $v_\ell(\cdot,t)$ is computed on $\mathcal{T}_\ell$ assuming (4.52), i.e. that the time steps Δt_ℓ are chosen subject to the CFL constraint (4.16).

The moments $\mathcal{M}^k(u)(t_1,...,t_k)$ of order $k \ge 2$ (resp. the k-th order space-time correlation functions) in (3.35) of the random entropy solution u can be estimated in the same way: based on (4.40) in Definition 4.5, the straightforward generalization along the lines of the MLMC estimate (4.54) of the MC estimate (4.41) for $\mathcal{M}^k(u)(t)$ leads to the definition of the MLMC-FVM estimator

$$E^{L,(k)}[u(\cdot,t)] := \sum_{\ell=0}^{L} E_{M_\ell}[(v_\ell(\cdot,t))^{(k)} - (v_{\ell-1}(\cdot,t))^{(k)}] , \quad 0 < t < \infty . \tag{4.55}$$

This generalizes (4.54) to moments $\mathcal{M}^k(u)(t)$ of order $k > 1$ [1]

[1]We assume here for notational convenience that $t_1 = t_2 = ... = t_k = t$. This implies that our k-th moment estimate only requires access to the FVM solutions at time t. The following developments directly generalize to the analysis of k-point temporal correlation functions of the random entropy solution as well; in this case, however, access to the full history of FVM solutions $v_\ell(\cdot,t)$ for $0 \le t \le T < \infty$ is required.

4.4.3. *Convergence Analysis.* We first analyze the MLMC-FVM mean field error

$$\left\|\mathbb{E}[u(\cdot,t)] - E^L[u(\cdot,t)]\right\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \tag{4.56}$$

for $0 < t < \infty$ and $L \in \mathbb{N}$. In particular, we are interested in the choice of the sample sizes $\{M_\ell\}_{\ell=0}^\infty$ such that, for every $L \in \mathbb{N}$, the MLMC error (4.56) is of order $(\Delta t_L)^s$, where s is the order of convergence in the Kusznetsov type error bound (4.32). The principal issue in the design of MLMC-FVM is the optimal choice of $\{M_\ell\}_{\ell=0}^\infty$ such that, for each L, an error (4.56) is achieved with minimal total work given by (based on (4.45))

$$\mathrm{Work}_L = \sum_{\ell=0}^{L} M_\ell O(\Delta x_\ell^{-d-1}) = O\left(\sum_{\ell=0}^{L} M_\ell \Delta x_\ell^{-d-1}\right) . \tag{4.57}$$

To estimate (4.56), we write (recall that $v_{-1} := 0$) using the triangle inequality, the linearity of the mathematical expectation $\mathbb{E}[\cdot]$ and the definition (4.54) of the MLMC estimator

$$\begin{aligned}
&\|\mathbb{E}[u(\cdot,t)] - E^L[u(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\
\leq\ & \|\mathbb{E}[u(\cdot,t)] - \mathbb{E}[v_L(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} + \left\|\mathbb{E}[v_L(\cdot,t)] - E^L[u(\cdot,t)]\right\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\
=\ & \|\mathbb{E}[u(\cdot,t)] - \mathbb{E}[v_L(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} + \left\|\sum_{\ell=0}^{L} \mathbb{E}[v_\ell - v_{\ell-1}] - E_{M_\ell}[v_\ell - v_{\ell-1}]\right\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\
=:\ & \quad \mathrm{I} \quad + \quad \mathrm{II} \quad .
\end{aligned}$$

We estimate terms I and II separately. By linearity of the expection, Term I equals

$$I = \|\mathbb{E}[u(\cdot,t) - v_L(\cdot,t)]\|_{L^1(\mathbb{R}^d)} = \|u(\cdot,t) - v_L(\cdot,t)\|_{L^1(\Omega;L^1(\mathbb{R}^d))}$$

which can be bounded by (4.36) with $k = 1$ and with the approximation property (4.25). We hence focus on term II and estimate further as follows:

$$\begin{aligned}
\mathrm{II} \ &\leq \ \sum_{\ell=0}^{L} \|\mathbb{E}[(v_\ell - v_{\ell-1})(\cdot,t)] - E_{M_\ell}[(v_\ell - v_{\ell-1})(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\
&\overset{(4.6)}{\leq} \ \sum_{\ell=0}^{L} M_\ell^{-\frac{1}{2}} \Big(\int_\Omega \|v_\ell(\cdot,t;\omega) - v_{\ell-1}(\cdot,t;\omega)\|_{L^1(\mathbb{R}^d)}^2 \, d\mathbb{P}(\omega)\Big)^{\frac{1}{2}} \\
&= \ \sum_{\ell=0}^{L} M_\ell^{-\frac{1}{2}} \|v_\ell(\cdot,t) - v_{\ell-1}(\cdot,t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \, .
\end{aligned}$$

We estimate for every $\ell \geq 0$ the size of the detail $v_\ell - v_{\ell-1}$ with the triangle inequality

$$\begin{aligned}
&\|v_\ell(\cdot,t) - v_{\ell-1}(\cdot,t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\
&\leq \|u(\cdot,t) - v_\ell(\cdot,t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))} + \|u(\cdot,t) - v_{\ell-1}(\cdot,t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \, .
\end{aligned}$$

Using here (4.36) with $\bar{t} = t$, $k = 2$ and (4.52), we obtain for every $\ell \in \mathbb{N}$ the estimate

$$\begin{aligned}
&\|(v_\ell - v_{\ell-1})(\cdot,t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \\
&\quad \leq \|u_0 - v_\ell^0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} + \|u_0 - v_{\ell-1}^0\|_{L^2(\Omega;L^1(\mathbb{R}^d))} + Ct\Delta x_\ell^s \, \|TV(u_0)\|_{L^2(\Omega;d\mathbb{P})} \, .
\end{aligned}$$

Using that for $0 \le s \le 1$, the cell-averages v_ℓ^0 satisfy, for every $k \in \mathbb{N}$ and for every $1 \le q \le \infty$,

$$\|u_0 - v_\ell^0\|_{L^k(\Omega;L^q(\mathbb{R}^d))} \le C\Delta x_\ell^s \|u_0\|_{L^k(\Omega;W^{s,q}(\mathbb{R}^d))}\,,$$

we arrive at the error bound

$$\|(v_\ell(\cdot,t)-v_{\ell-1})(\cdot,t)\|_{L^2(\Omega;L^1(\mathbb{R}^d))} \le C\big\{t\,\|TV(u_0)\|_{L^2(\Omega;d\mathbb{P})}+\Delta x_\ell^s\,\|u_0\|_{L^2(\Omega;W^{s,1}(\mathbb{R}^d))}\big\}\,.$$

Summing this error bound over all discretization levels $\ell = 0, ..., L$, we have proved the main result of the present paper.

Theorem 4.8. *Assume* (3.1) - (3.4), (3.13) - (3.16) *and* (4.50) - (4.52). *Then, for any sequence* $\{M_\ell\}_{\ell=0}^\infty$ *of sample sizes at mesh level* ℓ, *we have for the MLMC-FVM estimate* $E^L[u(\cdot,t)]$ *in* (4.54) *the error bound*

$$\begin{aligned}
&\|\mathbb{E}[u(\cdot,t)] - E^L[u(\cdot,t)]\|_{L^2(\Omega;L^1(\mathbb{R}^d))}\\
&\le C\big\{\bar{t}\,\Delta x_L^s\|TV(u_0)\|_{L^1(\Omega,d\mathbb{P})} + \Delta x_L^s\|u_0\|_{L^\infty(\Omega;W^{s,1}(\mathbb{R}^d))}\big\}\\
&\quad +C\Big\{\sum_{\ell=0}^L M_\ell^{-\frac{1}{2}}\Delta x_\ell^s\Big\}\big\{\|u_0\|_{L^2(\Omega;W^{s,1}(\mathbb{R}^d))} + t\,\|TV(u_0)\|_{L^2(\Omega;d\mathbb{P})}\big\}\,.
\end{aligned} \tag{4.58}$$

Theorem 4.8 is the basis for an optimization of the numbers M_ℓ of MC samples across the mesh levels. Our optimization of the level dependent selection of the Monte Carlo sample sizes M_ℓ will be based on the last term in the error bound (4.58): we select in (4.58) the M_ℓ such that as $\Delta t \downarrow 0$, all terms equal the Kusznetsov bound Δt_L^s in (4.32) at the finest level L. This implies

$$M_\ell^{-\frac{1}{2}}\,\Delta x_l^s \stackrel{!}{=} \hat{C}\Delta x_L^s,\quad \ell = 0,\dots,L-1\,. \tag{4.59}$$

Here, $\hat{C}$ is some integer constant that is independent of l, L. Using that

$$\Delta x_\ell = O(2^{-\ell}),\quad \ell = 0,1,2,\dots$$

we find $M_\ell = \hat{C}\Delta x_\ell^{2s}\Delta x_L^{-2s} = O(2^{2(L-\ell)s})$. This implies in (4.58) the bound

$$\|\mathbb{E}[u(\cdot,t)] - E^L[u(\cdot,t)]\|_{L^1(\mathbb{R}^d)} \le C(t)\Delta x_L^s\,, \tag{4.60}$$

while the total cost is, by (4.57), bounded by

(4.61)

$$\begin{aligned}
\mathrm{Work}_L &= \sum_{\ell=0}^L M_\ell\, O(\Delta x_\ell^{-(d+1)}) \le C\sum_{\ell=0}^L 2^{2(L-\ell)s+\ell(d+1)}\\
&= C\,2^{2Ls}\sum_{\ell=0}^L 2^{(d+1-2s)\ell} = C\,2^{2Ls+[(d+1)-2s]L} = C\,2^{(d+1)L} = O(\Delta x_L^{-(d+1)})
\end{aligned}$$

provided that the order s of the FVM satisfies

$$0 \le s < (d+1)/2\,. \tag{4.62}$$

We compare this bound with the work for a single level MC-FVM: by (4.45), we have the error bound (4.13) with total work

$$\mathrm{Work}(M_L;\mathcal{T}_L) = O(M_L\,\Delta x_L^{-(d+1)}) \stackrel{(4.46)}{=} O(\Delta x_L^{-(d+1+2s)})\,. \tag{4.63}$$

Inserting (4.61) into the asymptotic error bound (4.60), we obtain the following error estimate in terms of work

$$\|\mathbb{E}[u(\cdot,t)] - E^L[u(\cdot,t)]\|_{L^1(\mathbb{R}^d)} \leq C(\mathrm{Work}(M_L;\mathcal{T}_L))^{-s/(d+1)}\,. \tag{4.64}$$

We observe that *under the provision (4.62) the MLMC-FVM (4.61) behaves, in terms of accuracy versus work, as $L\to\infty$, exactly as the deterministic FVM* where the error vs. work was estimated in (4.33). We also observe that *for high order schemes, specifically when the condition* (4.62) *is violated, i.e. when $2s \geq d+1$, the complexity of the MLMC-FVM increases above that of one deterministic FVM solve of the same order, but is superior compared to the standard MCM.* In this case, more sophisticated "polynomial chaos" type discretizations with respect to the stochastic variable ω are required (see, e.g. [20] for details on this) to achieve optimal efficiency.

4.5. **Sparse Tensor MLMC-FVM.** The work to form a single tensor product in the estimates (4.41), (4.55) over a finite spatial domain $D\subset\mathbb{R}^d$ grows 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-th moment estimation, we introduce similar to strategy for high order moment approximation in elliptic problems with random data in [23, 2]

4.5.1. *Sparse Tensorization of FV Solutions.* Since the linear mappings $P_\ell : L^1(\mathbb{R}^d)\to S_\ell$ defined in (4.51), (4.23) are onto, we may define the linear space of *increment* or *details* of FV functions between successive meshes in the grid hierarchy $\mathfrak{M}=\{\mathcal{T}_\ell\}_{\ell=0}^\infty$ by

$$W_\ell := (P_\ell - P_{\ell-1})S_\ell, \qquad \ell\geq 0 \tag{4.65}$$

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

$$S_L = W_0\oplus ... \oplus W_L = \bigoplus_{\ell=0}^{L} W_\ell \tag{4.66}$$

and the k-point correlation functions $(v_L(\cdot,t))^{(k)}$ of the FV solutions on mesh $\mathcal{T}_L$ at time $t>0$ take values in the tensor product space

$$(S_L)^{(k)} := \underbrace{S_L\otimes ...\otimes S_L}_{k \text{ times}} = \sum_{|\underline{\ell}|_\infty\leq L} S_{\ell_1}\otimes ...\otimes S_{\ell_k} = \bigoplus_{|\underline{\ell}|_\infty\leq L}\bigotimes_{j=1}^{k} W_{\ell_j}\,. \tag{4.67}$$

Then, the full tensor projections

$$P_L^{(k)} v := \underbrace{P_L\otimes ...\otimes P_L}_{k \text{ times}} : L^1(\mathbb{R}^{kd})\to (S_L)^{(k)} \tag{4.68}$$

are bounded, linear and onto. Here, $|\underline{\ell}|_\infty := \max\{\ell_1,...,\ell_k\}$ and the last sum in (4.67) is a direct one. Obviously, if $N_L := \dim S_L < \infty$ (as is the case when, e.g., the spaces S_ℓ are only defined on a bounded domain $D\subset\mathbb{R}^d$) then $\dim((S_L)^{(k)}) = N_L^k$ which is prohibitive. Sparse Tensor approximations of k-point correlations $(v(\cdot,t))^{(k)}$ will be approximations in tensor products of spaces of piecewise constant functions on meshes on coarser levels which are defined similar to (4.67) by

$$\widehat{(S_L)}^{(k)} := \bigoplus_{|\underline{\ell}|_1\leq L}\bigotimes_{j=1}^{k} W_{\ell_j} \tag{4.69}$$

where now $|\underline{\ell}|_1 := \ell_1 + ... + \ell_k$. If the mesh family $\mathfrak{M}$ 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

$$\dim(\widehat{S_L})^{(k)} = O(N_L(\log_2 N_L)^{k-1}) \,. \tag{4.70}$$

With $(\widehat{S_L})^{(k)}$ in (4.69), we also define the *sparse tensor projection*

$$\widehat{(P_L)}^{(k)} := \bigoplus_{|\underline{\ell}|_1 \le L} \bigotimes_{j=1}^{k} (P_{\ell_j} - P_{\ell_j - 1}) : L^1(\mathbb{R}^{kd}) \to (\widehat{S_L})^{(k)} \,. \tag{4.71}$$

The approximation properties of the sparse tensor projection are as follows (cf. the Appendix): for any function $U(x_1, ..., x_k)$ which belongs to $(W^{s,1}(\mathbb{R}^d))^{(k)}$, it holds

$$\|U - \widehat{(P_L)}^{(k)} U\|_{L^1(\mathbb{R}^{kd})} \le C(\Delta x_L)^s |\log \Delta x_L|^{k-1} \|U\|_{(W^{s,1}(\mathbb{R}^d))^{(k)}} \tag{4.72}$$

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

4.5.2. *Definition of the Sparse Tensor MLMC-FVM Estimate.* With the above notions in hand, we proceed to the *definition of the sparse tensor MLMC-FVM estimator of* $\mathcal{M}^{(k)}(u(\cdot,t))$. To this end, we modify the full tensor product MLMC estimator (4.55) as follows (recall from (4.39) that $E_M[\cdot]$ denotes the MC estimate based on M samples): for a given sequence $\{M_\ell\}_{\ell=0}^L$ of MC samples at level ℓ, the sparse tensor MLMC estimate of $\mathcal{M}^k[u(\cdot,t)]$ is, for $0 < t < \infty$, defined by

$$\widehat{E^{L,(k)}}[u(\cdot,t)] := \sum_{\ell=0}^{L} E_{M_\ell}[\widehat{P_\ell}^{(k)}(v_\ell(\cdot,t))^{(k)} - \widehat{P_{\ell-1}}^{(k)}(v_{\ell-1}(\cdot,t))^{(k)}] \,. \tag{4.73}$$

We observe that (4.73) is identical to (4.55) 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 our convergence analysis ahead will show, use of sparse rather than full tensor products will not entail any reduction in the order of convergence of the k-th moment estimates.

4.5.3. *Error and Complexity Analysis of the sparse tensor MLMC-FVM.* We now generalize Theorems 4.2, 4.6 and 4.8.

Theorem 4.9. *Assume that (4.42) and that* (3.13) - (3.15) *hold. Assume further that we are given a FVM* (4.17) - (4.20) *such that* (4.16) *holds and such that Assumption* 4.3 *is satisfied; in particular, assume that the deterministic FVM scheme converges at rate* $s > 0$ *in* $L^\infty([0,\infty]; L^1(\mathbb{R}^d))$. *Then the MLMC-FVM estimate*

$\widehat{E^{L,(k)}}[u(\cdot,t)]$ defined in (4.73) satisfies, for every sequence $\{M_\ell\}_{\ell=0}^L$ of MC samples, the error bound

$$\|\mathcal{M}^k u(\cdot,t) - \widehat{E^{L,(k)}}[u(\cdot,t;\omega)]\|_{L^2(\Omega;L^1(\mathbb{R}^{kd}))}$$

$$\lesssim (1\vee t)\Delta x_L^s |\log \Delta x_L|^{k-1}\left\{\|\mathrm{TV}(u_0(\cdot,\omega))\|^k_{L^k(\Omega;d\mathbb{P})} + \|u_0(\cdot\,;\omega)\|^k_{L^\infty(\Omega;W^{s,1}(\mathbb{R}^d))}\right\}$$

$$+\left\{\sum_{\ell=0}^L \frac{\Delta x_\ell^s |\log \Delta x_\ell|^{k-1}}{M_\ell^{1/2}}\right\}\left\{\|u_0(\cdot\,;\omega)\|^k_{L^{2k}(\Omega;W^{s,1}(\mathbb{R}^d))} + t\|\mathrm{TV}(u_0(\cdot\,;\omega))\|^k_{L^{2k}(\Omega;d\mathbb{P})}\right\}\,.$$

The total work to compute the MLMC estimates $\widehat{E^{L,(k)}}[u(\cdot\,;t)]$ on compact domains $D\subset\mathbb{R}^d$ is therefore (with $O(\cdot)$ depending on the size of D)

$$\widehat{\mathrm{Work}}_L^{MLMC} = O\left(\sum_{\ell=0}^L M_\ell \Delta x_\ell^{-(d+1)} |\log \Delta x|^{k-1}\right)\,. \tag{4.74}$$

Based on Theorem 4.9, we infer that the choice (4.59) of numbers M_ℓ of MC samples at level ℓ should also be used in the MLMC estimation of moments of order $k>1$ of the random entropy solution, provided the order s of the underlying deterministic FVM scheme (4.17) - (4.19) satisfies the bound (4.62). The conversion of the FVM approximations of the draws $\hat{u}^i(\cdot\,,t;\omega)$ of the random solution at time $t>0$ into a multilevel basis and the sparse tensor product formation in the MLMC estimator (4.73) increases the work bounds (4.63) for the first moments only by a logarithmic factor, so that, in terms of the computational work, we have with the choices (4.59) of MC samples M_ℓ, the following error bound in terms of work in a bounded domain $D\subset\subset\mathbb{R}^d$:

$$\|\mathcal{M}^k u(\cdot,t) - \widehat{E^{L,(k)}}[u(\cdot,t;\omega)]\|_{L^2(\Omega;L^1(D^k))} \le C(\widehat{\mathrm{Work}}_L^{MLMC})^{-s'/(d+1)} \tag{4.75}$$

for any $0<s'<s$ with the constant depending on D and growing as $0<s'\to s\le 1$.

5. Numerical Examples

In this section, we present several numerical experiments to compare the standard MC-FVM and the MLMC-FVM. Both schemes are based on a underlying finite volume deterministic solver. The aim is to compare the performance of both schemes and corroborate the analysis presented in the previous section.

5.1. **Burgers' equation.**

We consider Burgers' equation,

$$u_t(x,t,\omega) + \left(\frac{u^2(x,t,\omega)}{2}\right)_x = 0, \quad x\in[0,1], \quad t>0, \text{and} \quad \omega\in\Omega\,. \tag{5.1}$$

For simplicity, we discretize the computational domain uniformly in space and use a monotone numerical scheme,

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}\left(F(u_j^n,u_{j+1}^n) - F(u_{j-1}^n,u_j^n)\right). \tag{5.2}$$

Here, Δt and Δx are the time step and mesh size respectively and are related to each other by the CFL condition (4.16). We denote

$$u_j^n \approx \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,t^n)dx$$

and the numerical flux was chosen as the Rusanov flux,

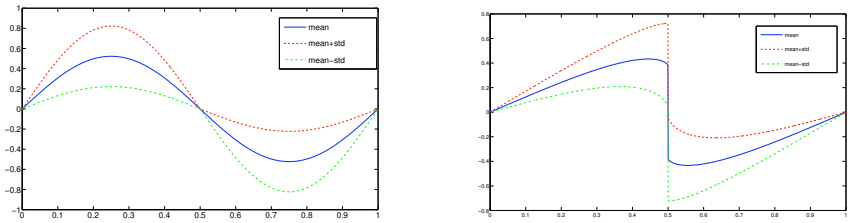
$$F(u_j^n, u_{j+1}^n) = \frac{1}{2}\left(f(u_j^n) + f(u_{j+1}^n) - \max\{|f'(u_j^n)|, |f'(u_{j+1}^n)|\}(u_{j+1}^n - u_j^n)\right), \tag{5.3}$$

with flux $f(u) = \frac{1}{2}u^2$. Note that the Rusanov flux results in the scheme (5.2) being monotone, consistent and conservative, [9]. Hence the above scheme satisfies the conditions of Assumption 4.3 and converges with rate 1/2 in the deterministic case. Transparent Neumann type boundary conditions (i.e, $\partial_\nu u = 0$ on the boundary ∂D of the computational domain, with ν being the unit outward normal) were used in all numerical experiments.

5.2. **Initial data with uncertain amplitude.** In this experiment, we consider the random Burgers' equation with parametric initial data,

$$u_0(x, \omega) = Y(\omega)\sin(2\pi x). \tag{5.4}$$

Here, $Y(\omega)$ is a uniformly distributed random variable taking values in $(0, 1)$ i.e, $Y \sim \mathcal{U}(0, 1)$. The Burgers' equation with above initial data is solved with both the MC-FVM and MLMC-FVMs based on the deterministic scheme (5.2). In order to assess discretization and sampling errors, a numerical reference solution was computed with $M = 10000$ samples on a uniform mesh of $2^{12} = 4096$ points. The initial conditions and the reference solution at time $t = 0.4$ are shown in Figure 1. As shown in Figure 1, the initial data is smooth but uncertain. As expected,



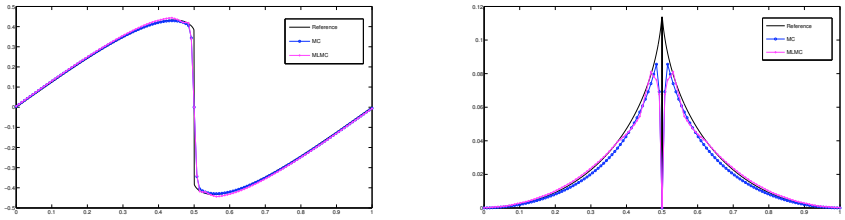
(a) t=0 (b) t=0.4

FIGURE 1. Reference solution for the stochastic Burgers' equation (5.1) with uncertain initial amplitude (5.4). Color graphs depict mean and mean ± standard deviation.

the smooth data evolves into discontinuities in the physical space and a shock has formed near $x = 0.5$ at time $t = 0.4$. We compute the mean reference solution and its standard deviation from the estimated second moment of the reference solution. As is apparent from Figure 1, the random entropy solution's mean field is also discontinuous in space.

We implemented the MC-FVM and the MLMC-FVM. Results for both methods are shown in Figure 2. Both schemes are compared on a finest mesh of $N_L = 128$ points. For the MC-FVM, we choose $M = 128$ MC samples as the formula $M = (\Delta x)^{-2s}$ in (4.46) reduces to $M = N$ (N being the number of mesh points) since the convergence rate is $s = \frac{1}{2}$ for the "first-order" scheme (5.2). The MLMC-FVM is based on $L = 5$ consecutive levels with maximum level consisting of $N_L = 2^7 = 128$

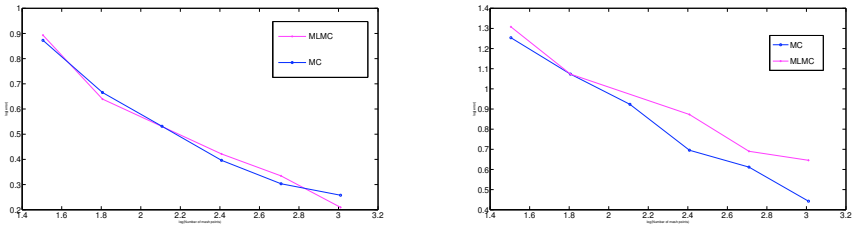
cells and minimum level L_0 consisting of $N_0 = 2^3$ mesh points. The number M_l of MC samples at the mesh $\mathcal{T}_l$ are chosen according to formula (4.59) realized in this particular case as $M_l = 2^{L-l}M_L$ with L being the maximum resolution. We choose $M_L = 8$ in this case. The mean and variance of the random solution are computed from the corresponding estimators and shown in Figure 2. Figure 2 shows that solutions computed with MC and MLMC schemes are comparable. The MC solutions appears to be slightly more accurate for both the mean and the variance. In order to quantify the errors, we compute the relative error in the mean and in the



(a) mean (b) variance

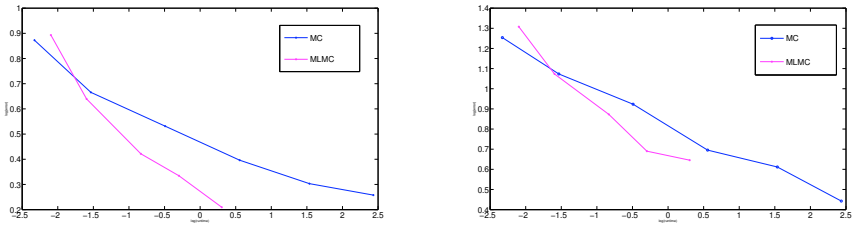
FIGURE 2. Computed solutions for the stochastic Burgers' equation (5.1) with uncertain initial amplitude (5.4). Color graphs depict mean and the variance at time $t = 0.4$ with the MC-FVM and MLMC-FVM.

variance (details of how this error is computed will be described later in this section) and present the error vs. mesh resolution in Figure 3. The results are plotted in a log-log format with relative error in the y-axis and the number of mesh points in the x-axis. The number of samples at each resolution is fixed by the formulas described before. In particular, we take 8 different levels for the MLMC scheme with the finest resolution consisting of 2^{10} mesh points. The number of samples for the finest resolution of the MLMC-FVM is fixed at $M_L = 8$. The results show that both, MC-FVM and MLMC-FVMs converge in the mean and variance at a rate slightly better than the expected rate of 0.5. The differences in error between both methods is minor for the mean whereas the MC-FVM has consistently lower errors in the variance. The analysis in the previous section suggests that the MC-FVM and the MLMC-FVM will be comparable in accuracy at the same resolution of physical space. The principal difference between the two methods lies in their efficiency, measured in terms of work or computational time. Hence, we measure the runtime (in seconds) for each scheme and present the error vs. runtime in Figure 4. The results are plotted in log-log and show that there is a consistent gain in efficiency with the MLMC-FVM. For the mean, the speedup achieved with the MLMC-FVM is of at least two orders of magnitude (a factor of about 100) as compared to the MC-FVM. This considerable gain in efficiency achieved by MLMC-FVM allows accurate numerical solution of realistic hyperbolic conservation laws in several spatial dimensions with random initial data which were hitherto beyond the scope of numerical schemes. The speedup with respect to errors in variance is more modest but still at least an order of magnitude, on this problem.



(a) mean (b) variance

FIGURE 3. A log-log plot for the relative error (y-axis) vs. resolution (x-axis) for the stochastic Burgers' equation (5.1) with uncertain initial amplitude (5.4). MC-FVM and MLMC-FVM are compared.



(a) mean (b) variance

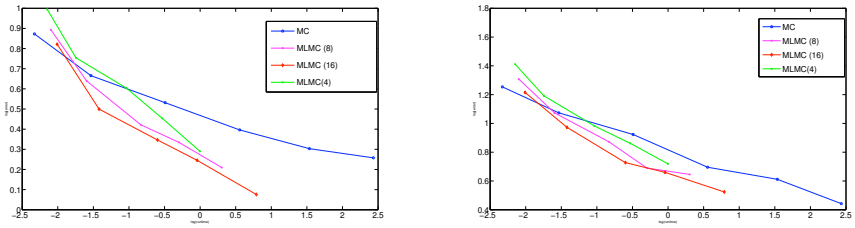
FIGURE 4. A log-log plot for the relative error (y-axis) vs. runtime (x-axis) for the stochastic Burgers' equation (5.1) with uncertain initial amplitude (5.4). MC-FVM and MLMC-FVM are compared.

5.3. **Effect of minimum number of samples on MLMC-FVM.** One of the free parameters in the above simulation was the choice of the number M_L of samples at the finest resolution of the MLMC-FVM. By the formula $M_l = M_L 2^{L-l}$, we see that the number of samples at the finest resolution might have a significant bearing on the results. In order to investigate this issue, we choose three different values of M_L i.e, $M_L = 4, 8, 16$ and present the relative error vs. runtime results in Figure 5. The results show that the speedup is at least asymptotically independent of M_L. For the computed resolutions, $M_L = 16$ appears to be a good choice as the speedup with respect to mean error is a factor of 130 and with respect of variance is a factor of 20. However, the choice $M_L = 4$ appears to be the most efficient asymptotically implying that very few samples need to be chosen at the finest resolution.

5.4. **Initial data with uncertain phase.** In this experiment, we consider the random Burgers' equation with the parametric random initial data,

$$u_0(x,\omega) = \sin(2\pi(x + 0.1Y(\omega)). \tag{5.5}$$

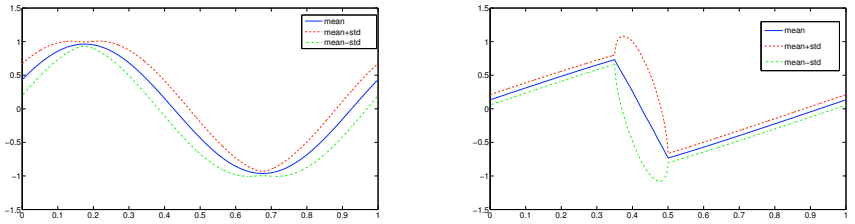
Here, $Y(\omega) \sim \mathcal{U}(0,1)$. The Burgers' equation with above initial data is solved with both the MC and MLMC methods based on the deterministic scheme (5.2). In order to compute errors, we computed a reference solution with $M = 10000$



(a) mean (b) variance

FIGURE 5. A log-log plot for the relative error (y-axis) vs. runtime (x-axis) for the stochastic Burgers' equation (5.1) with uncertain initial amplitude (5.4). The MC-FVM and MLMC-FVMs are compared and the sensitivity of the results to M_L, number of samples at the finest resolution of the MLMC methods is evaluated. We choose $M_L = 4, 8, 16$ in the above simulations.

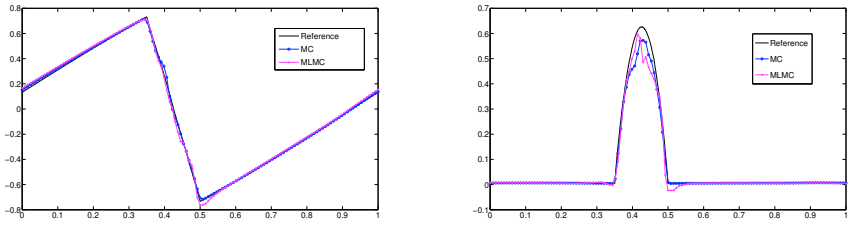
samples on a refined mesh of $2^{12} = 4096$ points. The initial conditions and the reference solution at time $t = 0.4$ are shown in Figure 6. As shown in Figure



(a) t=0 (b) t=0.4

FIGURE 6. Reference solution for the stochastic Burgers' equation (5.1) with uncertain initial phase (5.5). Color graphs depict mean and mean ± standard deviation.

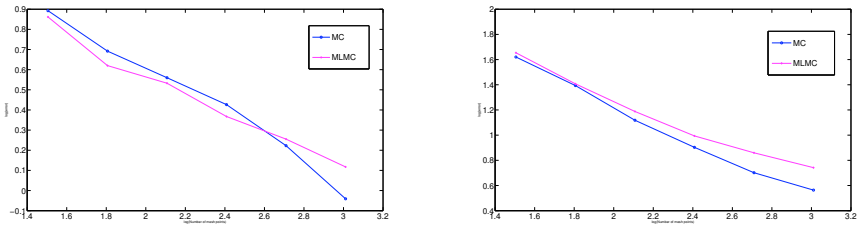
6, the mean field in this case is no longer discontinuous (compare with Figure 1) but is Lipschitz continuous. This rather surprising *smoothness* is generic for problems with uncertain shock locations (as in this case) and will be explained in a forthcoming paper. Furthermore, the variance in this case is concentrated at the shocks. We show results computed with the MC and MLMC methods for a resolution of 128 mesh points (5 different levels in the MLMC method) in analogy with the previous experiment. The results of Figure 7 show that both the MC and MLMC methods approximate the reference solution reasonably well. We also note that while the "pathwise" random entropy solutions are well known to develop a discontinuity in finite time, the ensemble average of the random entropy solution is continuous (we refer to [21]) for a more detailed analysis). In this case, the error of the MLMC-FVM appears to be larger than that of the MC-FVM. The quantitative comparison of both methods is shown in Figures 8 and 9. In Figure 8, we show a



(a) mean (b) variance

FIGURE 7. Computed solutions for the stochastic Burgers' equation (5.1) with uncertain initial phase (5.5). Color graphs depict mean and the variance at time $t = 0.4$ with the MC and MLMC methods.

log-log plot for the relative error (in both mean and variance) vis a vis the number of mesh points. The MLMC method uses $M_L = 16$ as the minimum number of samples. The results show that the expected rates of convergence are realized. For a fixed mesh resolution, the MC method is more accurate than the MLMC method. However, the MLMC method is much faster as shown in Figure 9 where the error vs. runtime is plotted. As before, the MLMC method displays a speedup of two orders of magnitude with respect to the mean and an order of magnitude with respect to the variance. As in the previous experiment, the asymptotic results were not sensitive to the number of samples at the finest resolution of the MLMC method.



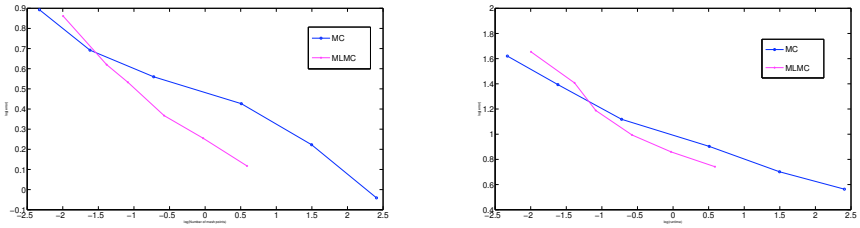
(a) mean (b) variance

FIGURE 8. A log-log plot for the relative error (y-axis) vs. resolution (x-axis) for the stochastic Burgers' equation (5.1) with uncertain initial phase (5.5). MC-FVM and MLMC-FVM are compared.

5.4.1. *Computation of the error.* The error estimates (4.43) for the MC-FVM and (4.58) for the MLMC-FVM are based on the norm $L^2(\Omega; L^1(\mathbb{R}^d))$. In order to compute the error in this norm, we fix an index k and compute the following relative error,

$$\mathcal{R}E_k^{m,v} = 100 \times \frac{\|U_{ref}^{m,v} - U_k^{m,v}\|_{l^1}}{\|U_{ref}^{m,v}\|_{l^1}}.$$

Here, m, v refer to the mean and the variance respectively and U_{ref} denotes the reference solution. The computed solution for index k is denoted by $U_k^{m,v}$. The



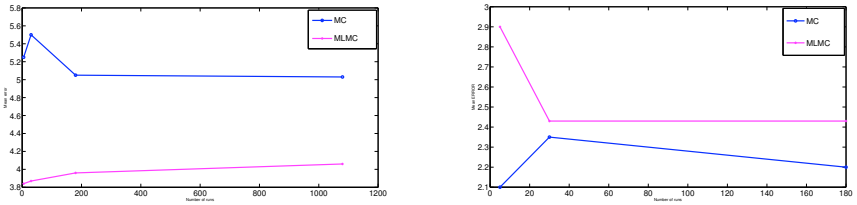
(a) mean (b) variance

FIGURE 9. A log-log plot for the relative error (y-axis) vs. runtime (x-axis) for the stochastic Burgers' equation (5.1) with uncertain initial phase (5.5). MC-FVM and MLMC-FVM are compared.

index k refers to independent multiple runs of the schemes and is varied in order to obtain different realizations of the probability space. Then, the error is summed over k according to

$$\mathcal{R}E^{m,v} = \sqrt{\sum_{k=1}^{K} (\mathcal{R}E_k^{m,v})^2}.$$

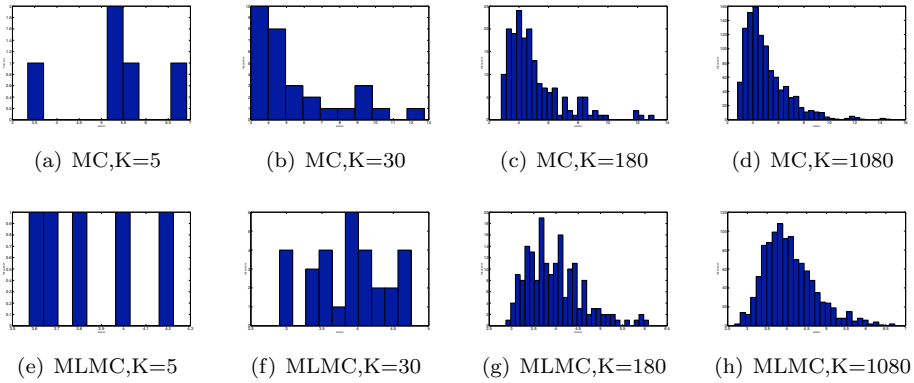
The free parameter at our disposal is the number K of independent runs for each



(a) 64pts (b) 256pts

FIGURE 10. The relative mean error vs. maximum run parameter K for the stochastic Burgers' equation (5.1) with uncertain initial phase (5.5). MC-FVM and MLMC-FVM are compared.

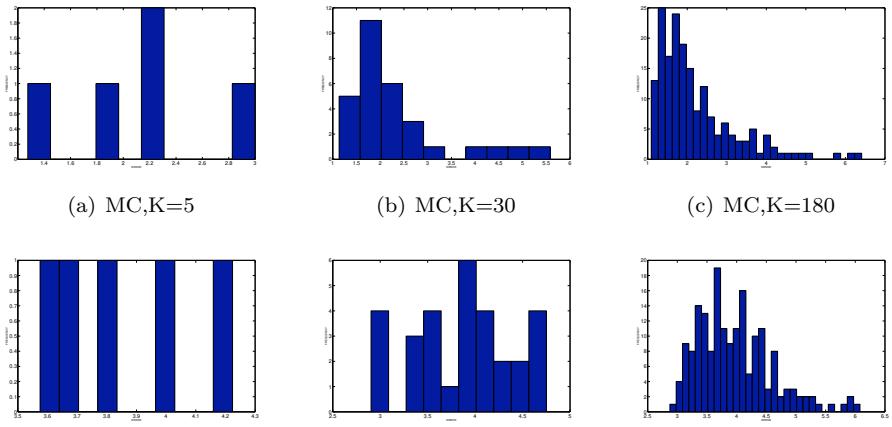
scheme. We investigate the sensitivity of the error with respect to this parameter in the following. In Figure 10, we plot the relative error (described above) for different values of K and for two different mesh resolutions consisting of 64 and 256 points respectively. We see from Figure 10 that there is some variation in the error for small values of K but they settle down to an approximate constant for moderate values of K ($K \approx 30$ in this example). Furthermore, the dependence with respect to K becomes even less pronounced when the mesh resolution is increased. The statistics for the error dependence with respect to the number of runs is shown by plotting empirical histograms in Figures 11 and 12 respectively. The results indicate that certain runs will lead to outliers in terms of either small but mostly large relative errors. Hence, one should use a moderate number of runs to compute the error, particularly on coarse mesh resolutions. Interestingly, the number of



(a) MC,K=5 (b) MC,K=30 (c) MC,K=180 (d) MC,K=1080

(e) MLMC,K=5 (f) MLMC,K=30 (g) MLMC,K=180 (h) MLMC,K=1080

FIGURE 11. Histograms showing the relative error in the mean for each run of the MC and MLMC method for different values of maximum run parameter $K = 5, 30, 180, 1080$. The mesh resolution is 64 points



(a) MC,K=5 (b) MC,K=30 (c) MC,K=180

(d) MLMC,K=5 (e) MLMC,K=30 (f) MLMC,K=180

FIGURE 12. Histograms showing the relative error in the mean for each run of the MC and MLMC method for different values of maximum run parameter $K = 5, 30, 180$. The mesh resolution is 256 points

outliers as well as their spread seems to be more pronounced for the MC method compared to the MLMC method.

5.5. **Euler equations.** The theory for the MC-FVM and MLMC-FVM has been presented in the case of scalar conservation laws in this paper. However, most of the interesting examples of conservation laws involve systems. We are constrained in our efforts to obtain rigorous convergence results for systems by the lack of rigorous error estimates or convergence results for the *deterministic* FVM for systems of

equations. However, we can readily extend the algorithms for systems and evaluate the performance of the MC-FVM and MLMC-FVM numerically. We will do so in this section.

Another limitation of the MLMC method as suggested by the error vs. work estimate in Section 4 is order condition (4.62). In one space dimension, this bound implies that the MLMC method will have the same error vs. work estimate as the deterministic finite volume scheme only if the convergence rate of the underlying finite volume scheme is less than 1. This condition is satisfied by first-order schemes and the numerical results in the preceding section corroborated the expected speedup of the MLMC method when compared to the MC method.

First-order schemes are rarely employed for practical computations as they are too diffusive. High resolution schemes based non-oscillatory limiter based reconstruction procedures are frequently used. We seek to investigate whether the MLMC method together with a high resolution formally second-order scheme will still be faster and more efficient than the standard MC method with the same underlying scheme. This evaluation is performed below.

We consider the Euler equations of gas dynamics in one space dimension,

$$\begin{aligned} \rho_t + (\rho u)_x &= 0, \\ (\rho u)_t + (\rho u^2 + p)_x &= 0, \\ E_t + ((E+p)u)_x &= 0. \end{aligned} \tag{5.6}$$

Here, ρ is the density, u is the velocity, p is pressure and E is the total energy. The variables are related by an ideal gas equation of state:

$$E = \frac{p}{\gamma - 1} + \frac{\rho u^2}{2},$$

with γ being the gas constant. The Euler equations are approximated by a standard first-order (in space and time) scheme of the form,

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \frac{\Delta t}{\Delta x}\left(\mathbf{F}(\mathbf{U}_{\mathbf{j}}^{\mathbf{n}}, \mathbf{U}_{j+1}^n) - \mathbf{F}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)\right). \tag{5.7}$$

Here, Δt and Δx are the time step and mesh size respectively for a uniform discretization. The vector of unknowns is denoted by $\mathbf{U} = \{\rho, \rho u, E\}$ and the flux vector by $\mathbf{f} = \{\rho u, \rho u^2 + p, (E+p)u\}$. For convenience, we consider the Rusanov flux:

(5.8)

$$\begin{aligned} \mathbf{F}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) \quad = \quad & \frac{1}{2}\left(\mathbf{f}(\mathbf{U}_j^n) + \mathbf{f}(\mathbf{U}_{j+1}^n)\right. \\ & \left. - \max\{|u_j^n| + c_j^n, |u_{j+1}^n| + c_{j+1}^n\}\mathbf{ID}_{3\times 3}(\mathbf{U}_{j+1}^n - \mathbf{U}_j^n)\right) . \end{aligned}$$

Here, c denotes the sound speed $c = \sqrt{\frac{\gamma p}{\rho}}$.

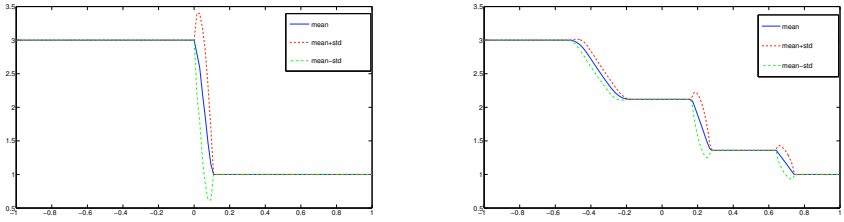
The scheme can be readily generalized to second-order of accuracy by using a piecewise linear reconstruction based on the minmod limiter and a strong stability preserving Runge-Kutta method, [17].

5.5.1. *Sod shock tube with uncertain shock location.* In this experiment, we consider the Euler equations (5.6) with Riemann initial data,

$$\{\rho_0(x,\omega), u_0(x,\omega), p_0(x,\omega)\} = \begin{cases} \{3.0, 0.0, 3.0\}, & \text{if} \quad x < 0.1Y(\omega), \\ \{1.0, 0.0, 1.0\}, & \text{if} \quad x > 0.1Y(\omega). \end{cases} \tag{5.9}$$

Here, $Y(\omega) \sim \mathcal{U}(0,1)$. Note that the above initial data is a random version of the standard Sod shock tube with an uncertain initial shock location.

The initial conditions and a reference solution (computed with a second-order MC scheme on a mesh of 2048 points with 10000 samples) are shown in Figure 13. The results show that the initial mean and variance break into three waves – one corresponding to the rarefaction wave, one to a contact and the right most wave to a shock in the deterministic case. However, the mean representations of the contact and shock are Lipschitz continuous. This is very similar to the results obtained for the scalar case (see Figure 6) where the mean field was also Lipschitz for the uncertain shock location case. In this case, the random solution's variance is concentrated at the contact discontinuity. We denote the MC scheme with a first-

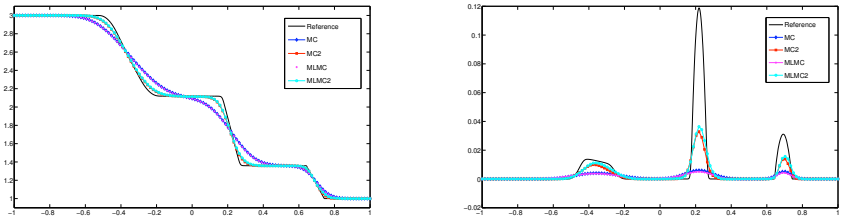


(a) t=0 (b) t=0.5

FIGURE 13. Reference solution for the stochastic Euler equations with uncertain initial shock location (5.9). Color graphs depict mean and mean ± standard deviation.

order FVM as MC and the MC scheme with a second-order FVM as the MC(2) scheme. The corresponding combination of the MLMC with first- and second-order FVM are termed as MLMC and MLMC(2), respectively. A qualitative comparison of all the four schemes is shown in Figure 14, where we compare the MC(2) and MLMC(2) schemes on a mesh of 128 points. The MLMC schemes include 5 levels of resolution with the finest resolution consisting of 128 points and the coarsest of 8 points. The number of samples for the MC and MC(2) scheme are set to 128, in analogy with the scalar case. For the MLMC scheme, the number of samples are chosen by the formula, $M_l = M_L 2^{L-l}$ with $M_L = 8$. The MLMC(2) scheme requires samples chosen by $M_l = M_L 2^{2(L-l)}$, based on the heuristic argument that the formally second-order schemes will converge twice as fast as the first-order scheme and the formula (4.59). The results suggest that the schemes are more diffusive for the Euler equations and the errors in mean but more acutely in the variance are larger than in the scalar case. Furthermore, the difference between the first-order and second-order schemes is substantial with the second-order schemes being considerably more accurate near the mean representations of shock and the contact. There are very minor difference between the MC and MLMC methods in this realization.

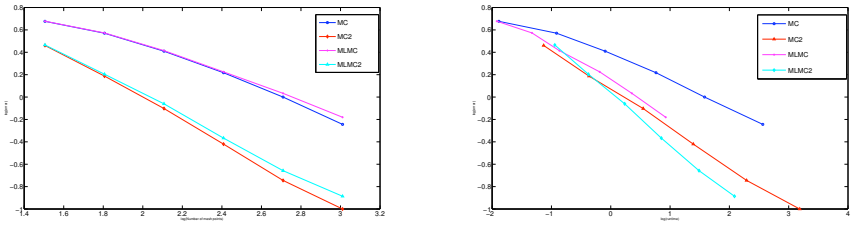
The errors in mean are quantified in Figure 15 where we plot the relative mean error vs. mesh resolution for all the four schemes. As expected, the second-order schemes are more accurate than the first-order schemes. Furthermore, the computed error is consistently lower with the MC schemes when compared to the MLMC



(a) mean (b) variance

FIGURE 14. Computed solutions for the stochastic Euler equations with uncertain initial shock location (5.9). Color graphs depict mean and the variance at time $t = 0.5$ with the MC(2) and MLMC(2) methods.

schemes, reinforcing the conclusions arrived in the scalar case. On the other hand, the MLMC schemes are faster as shown in the right graph of Figure 15 where we plot the error vs. runtime in log-log. The figure shows that the MLMC scheme is about two orders of magnitude faster (for the same relative error) as the MC scheme. This is similar to the speedup seen in the scalar case. More surprisingly, the MLMC(2) scheme is also about an order of magnitude faster than the MC(2) scheme. This is not predicted by the theory as the second-order case is not covered due to the bound (4.62).



(a) error vs. resolution (b) error vs. runtime

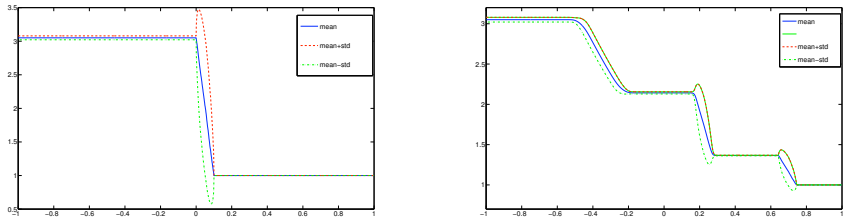
FIGURE 15. Computed solutions for the stochastic Euler equations with uncertain initial shock location (5.9). We show log of mean error at time $t = 0.5$ in the x-axis with respect of log of resolution and log of the runtime (y-axis) for both the MC(2) and MLMC(2) methods.

5.5.2. *Sod shock tube with uncertain shock location and uncertain amplitude.* In this experiment, we consider the Euler equations (5.6) with Riemann initial data,

(5.10)
$$\{\rho_0(x,\omega), u_0(x,\omega), p_0(x,\omega)\} = \begin{cases} \{3.0 + 0.1Z(\omega), 0.0, 3.0\}, & \text{if} \quad x < 0.1Y(\omega), \\ \{1.0, 0.0, 1.0\}, & \text{if} \quad x > 0.1Y(\omega). \end{cases}$$

Here, $Y(\omega), Z(\omega) \sim \mathcal{U}(0,1)$. Hence, we have uncertainty in terms of two parameters: uncertain shock location and uncertain amplitude.

The initial data and reference solution (computed with the same configuration as the previous numerical experiment) are shown in Figure 16. Comparing with

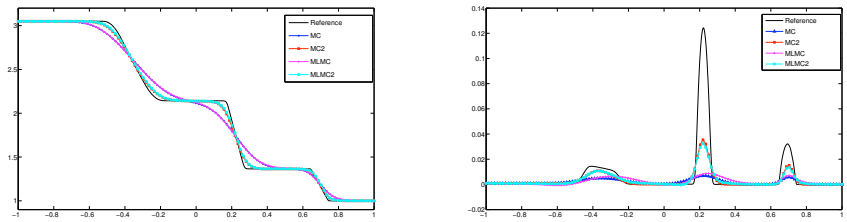


(a) t=0 (b) t=0.5

FIGURE 16. Reference solution for the stochastic Euler equations with uncertain initial shock location and amplitude (5.10). We show the mean and mean ± standard deviation.

the previous numerical experiment, we observe that there is a subtle competition between the two sources of uncertainty. There is some uncertainty in the amplitude of left most state but this uncertainty is considerably lower for the state to the left of the contact and almost vanishes for the state to the right of the contact. The nonlinearity appears to distribute uncertainty spatially with most of the uncertainty being in the location of the mean shock and mean contact. Furthermore, the mean shock and contact are both Lipschitz continuous.

We compare the four schemes on a mesh of 128 points in Figure 17 and find similar results to the previous experiment. The second-order schemes are considerably more accurate but for both first and second order schemes, there are very minor differences between the solutions computed with the MC and MLMC schemes.

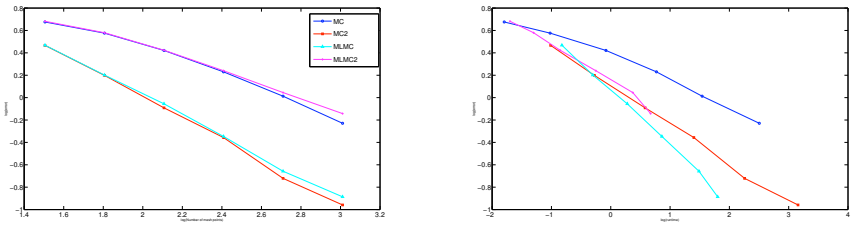


(a) mean (b) variance

FIGURE 17. Computed solutions for the stochastic Euler equations with uncertain initial shock location and amplitude (5.10). Color graphs depict mean and the variance at time $t = 0.5$ with the MC(2) and MLMC(2) methods.

The error vs. resolution and error vs. runtime plots shown in Figure 18 are very similar to those in the previous experiments. For a given resolution, the MC schemes

are more accurate than their MLMC counterparts at the same order. However, the MLMC schemes are faster. In particular, the MLMC scheme gives a speedup of two orders of magnitude compared to the MC scheme whereas the MLMC(2) scheme gives at least an order of magnitude compared to the MC(2) scheme. In fact, in this case the first-order MLMC scheme is more efficient than the second-order MC(2) scheme.



(a) error vs. resolution (b) error vs. runtime

FIGURE 18. Computed solutions for the stochastic Euler equations with uncertain initial shock location and amplitude (5.10). We show log of mean error at time $t = 0.5$ in the x-axis with respect of log of resolution and log of the runtime (y-axis) for both the MC(2) and MLMC(2) methods.

6. Conclusions

We consider scalar conservation laws in several space dimensions with uncertain initial data. The proper notion of random entropy solution is formulated and shown to be well-posed. Further, we show existence of higher moments of the solution and for the k-point correlation functions provided that the initial data has the desired regularity.

We propose Monte Carlo (MC) methods, together with standard finite volume schemes to approximate the random entropy solution. The MC-FVM is shown to converge to this solution but we derive accuracy vs. work estimates that indicate that the MC methods will be computationally slow.

Hence, we propose a new class of multi-level Monte Carlo (MLMC) methods and prove them to be convergent. These methods are designed to have the same accuracy vs. work estimate as a deterministic FVM (at least for low order schemes). Hence, these methods will be much faster than the standard MC-FVM.

We present several numerical experiments for scalar conservation laws in one space dimension that reinforce the theory. In particular, the MLMC method yields about two orders of magnitude speedup vis a vis the MC method in computing the mean. Furthermore, the speedup is more than an order of magnitude for computing the variance.

Although our theoretical results are restricted to scalar conservation laws, the MLMC algorithm can be easily extended to include systems like the Euler equations. Numerical experiments suggest that the MLMC method continues to yield a speedup of two order of magnitude over the MC method even in this case.

The theory places a limitation on the optimal complexity of the MLMC method by restricting it to low-order methods. In particular, the optimal rate is realized for first-order methods in one space dimension. However, computational results indicate that the MLMC method continues to yield a speedup of at least an order of magnitude, even when it is coupled with a second-order high resolution finite volume scheme.

The results of this paper clearly suggest replacing the standard MC method by the MLMC variant as it is much faster at comparable accuracy. Given a deterministic FV solver, the MLMC-FVM is nonintrusive and as easy to code and to parallelize as the standard MC-FVM. We therefore expect it to be competitive in terms of computational efficiency with respect to the widely used gPC stochastic Galerkin and collocation methods. A detailed comparison of the MLMC schemes with gPC discretizations will be considered in a forthcoming paper. Further work in progress includes implementing the MLMC-FVM in more than one space dimension. As the MLMC-FVM is obviously suited for massively parallel implementation, tackling realistic problems in two and three spatial dimensions with random data appears feasible.

7. Appendix: Analysis of sparse tensor approximations in L^1

We let $D \subset \mathbb{R}^d$ denote a bounded Lipschitz polyhedron with plane faces, and $\mathcal{T}_0 = \{K\}$ a regular triangulation of D into simplices K. For $\ell = 1, 2, ...$, we denote by $\mathcal{T}_\ell$ a sequence of triangulations obtained from $\mathcal{T}_0$ by uniform subdivision. Then $h_\ell = \max\{\mathrm{diam}(K) : K \in \mathcal{T}_\ell\} = 2^{-\ell} h_0$. For all $\ell \in \mathbb{N}_0$, each $K \in \mathcal{T}_\ell$ is an affine image of the unit simplex $\hat{K}$: for each $x \in K$, it holds $K \ni x = F_K(\hat{x}) = A_K \hat{x} + b_K$ with $\hat{x} \in \hat{K}$, and with $\det(A_K) = O(h_K^d)$. Recall that for $u \in L^1(D)$, $P_\ell u$ denotes the projection of u onto the piecewise constant functions on $\mathcal{T}_\ell$. We claim that there exists a constant $\hat{c}(s, \kappa)$ where κ is as in (4.50), such that for every $u \in W^{s,1}(D)$, $0 \le s \le 1$ holds

$$\forall \ell \in \mathbb{N}_0 \, , 0 \le s \le 1 : \quad \|u - P_\ell u\|_{L^1(D)} \le \hat{c}(s, \kappa) h_\ell^s |u|_{W^{s,1}(D)} \tag{7.1}$$

This estimate is a special case of general results on spline interpolation. We present here a direct proof of (7.1). We distinguish the cases $s = 0$, $s = 1$ and $0 < s < 1$. The case $s = 0$ is merely the $L^1(D)$-boundedness of the projector P_ℓ. In case $s = 1$, we note that in $\hat{K}$ holds a Poincaré-type inequality: there exists a constant $\hat{c} > 0$ such that for every $\hat{u} \in W^{1,1}(\hat{K})$

$$\|\hat{u} - \frac{1}{|\hat{K}|} \int_{\hat{K}} \hat{u} d\hat{x}\|_{L^1(\hat{K})} \le \hat{c} \|\hat{\nabla} \hat{u}\|_{L^1(\hat{K})} \, .$$

This follows by the standard argument that $\hat{c} = 0$ would contradict the compactness of the embedding $W^{1,1}(\hat{K}) \subset L^1(\hat{K})$ (see, e.g., [28, Theorem 4.2.1] with $m = 1$, $k = 0$ and $p = 1$). Mapping $\hat{K}$ affinely to $K \in \mathcal{T}_\ell$, we find

$$\forall K \in \mathcal{T}_\ell : \quad \|u - \frac{1}{|K|} \int_K u dx\|_{L^1(K)} \le c h_K \|\nabla u\|_{L^1(K)} \, .$$

Here, the constant $c > 0$ depends only on the shape regularity constant κ of $\mathcal{T}_\ell$. Summing the estimate over $K \in \mathcal{T}_\ell$ implies (7.1) for $s = 1$.

Assume next $0 < s < 1$ and recall the definition of the $|\circ|_{W^{s,1}(K)}$ seminorm on $K \in \mathcal{T}_\ell$:

$$|u|_{W^{s,1}(K)} := \int_{x\in K}\int_{y\in K} \frac{|u(x)-u(y)|}{|x-y|^{d+s}}dxdy.$$

From the (compact) embedding $W^{s,1}(\hat{K}) \subset L^1(\hat{K})$ we infer once more the existence of $\hat{c}(s) > 0$ such that

$$\|\hat{u} - \frac{1}{|\hat{K}|}\int_{\hat{K}} \hat{u}d\hat{x}\|_{L^1(\hat{K})} \leq \hat{c}(s)\int_{\hat{K}}\int_{\hat{K}} \frac{|\hat{u}(\hat{x})-\hat{u}(\hat{y})|}{|\hat{x}-\hat{y}|^{d+s}}d\hat{x}d\hat{y} = \hat{c}(s)|\hat{u}|_{W^{s,1}(\hat{K})}\ .$$

Inserting here the affine change of variables $x = A_K\hat{x} + b_K$, $y = A_K\hat{y} + b_K$ we find

$$\|u - \frac{1}{K}\int_K udx\|_{L^1(K)} \leq c(s,\kappa)h_K^s \int_K\int_K \frac{|u(x)-u(y)|}{|x-y|^{d+s}}dxdy = c(s,\kappa)h_K^s|u|_{W^{s,1}(K)}$$

for every $K \in \mathcal{T}_\ell$. Summing this bound over all $K \in \mathcal{T}_\ell$ implies for $u \in W^{s,1}(D)$

$$\|u - P_\ell u\|_{L^1(D)} \leq c(s,\kappa)\sum_{K\in\mathcal{T}_\ell} h_K^s|u|_{W^{s,1}(K)} \leq c(s,\kappa)h_\ell^s|u|_{W^{s,1}(D)}$$

which implies (7.1) for $0 < s < 1$. The error bound (4.72) follows then in a standard way (see, e.g. [25]).

References

[1] R. Abgrall. *A simple, flexible and generic deterministic approach to uncertainty quantification in non-linear problems*. Rapport de Recherche, INRIA, 2007.

[2] A. Barth, Ch. Schwab and N. Zollinger, *Multilevel MC Method for Elliptic PDEs with Stochastic Coefficients* , Report, SAM, (2010) (in review) .

[3] Q. Y. Chen, D. Gottlieb and J. S. Hesthaven. *Uncertainty analysis for steady flow in a dual throat nozzle.* J. Comput. Phys, 204, 2005, 378-398.

[4] Giuseppe Da Prato and Jerzy Zabcyk, *Stochastic Equations in infinite dimensions*, Cambridge Univ. Press (1991).

[5] Constantine M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics (2nd Ed.)*, Springer Verlag (2005).

[6] Weinan E, K. Khanin, A. Mazel and Y. Sinai. *I*nvariant measures for the Burgers equation with stochastic forcing. Ann. Math., 151, 2000, 877-960.

[7] M. Giles. *Improved multilevel Monte Carlo convergence using the milstein scheme.* Preprint NA-06/22, Oxford computing lab, Oxford, U.K, 2006.

[8] M. Giles. *Multilevel Monte Carlo path simulation.* Oper. Res., 56, 2008, 607-617.

[9] Edwige Godlewski and Pièrre A. Raviart, *Hyperbolic Systems of Conservation Laws* Mathematiques et Applications, Ellipses Publ. Paris (1991).

[10] Edwige Godlewski and Pièrre A. Raviart, *The numerical solution of multidimensional Hyperbolic Systems of Conservation Laws*, Springer Verlag Berlin Heidelberg New York (1995).

[11] Eymard, Robert and Gallouët, Thierry and Herbin, Raphaèle, *Finite volume methods*, in Handbook of numerical analysis, Vol. **VII**, pp. 713–1020, North-Holland, Amsterdam, (2000).

[12] S. Heinrich. *Multilevel Monte Carlo methods* in Large-scale scientific computing, Third international conference LSSC 2001, Sozopol, Bulgaria, 2001, Lecture Notes in Computer Science, Vol **2170**, Springer Verlag (2001), pp. 58-67.

[13] H. Holden and N.H. Risebro, *Front Tracking for Hyperbolic Conservation Laws* Applied Mathematical Sciences **152**, 2nd Printing Springer Verlag (2007).

[14] H. Holden and N. H. Risebro, *C*onservation laws with a random source, Applied Mathematics and Optimization. An International Journal with Applications to Stochastics **36**(1997)(2) pp. 229-241.

[15] H. Holden, T. Lindstrøm, B. Øksendal, J. Ubøe and T.-S. Zhang, *T*he Burgers equation with a noisy force and the stochastic heat equation, Comm. Partial Differential Equations **19**(1994)(1-2) pp. 119-141.

[16] D. Kröner, Numerical Schemes for Conservation Laws, J. Wiley and B.G. Teubner Publ., Chichester UK and Stuttgart, FRG (1997)
[17] R.A. LeVeque, *Numerical Solution of Hyperbolic Conservation Laws* Cambridge Univ. Press 2005.
[18] W.A. Light and E.W. Cheney, *Approximation Theory in Tensor Product Spaces* Lecture Notes in Mathematics **1169**, Springer Verlag Heidelberg, New York (1980).
[19] G. Lin, C.H. Su and G. E. Karniadakis. *The stochastic piston problem.* PNAS **101**(2004), pp. 15840-15845.
[20] S. Mishra and Ch. Schwab, *Entropy Stability of stochastic Galerkin and inifinite dimensional hyperbolic systems obtained from general moment closures* Working Paper SAM, 2010.
[21] S. Mishra and Ch. Schwab, *(in preparation)* Working Paper SAM, 2010.
[22] G. Poette, B . Després and D. Lucor. *Uncertainty quantification for systems of conservation laws.* J. Comput. Phys. **228**(2009) pp. 2443-2467.
[23] T. von Petersdorff and Ch. Schwab, *Sparse Finite Element Methods for Operator Equations with Stochastic Data*, Applications of Mathematics **51**(2) (2006) 145-180.
[24] J. Smoller, *Shock waves and reaction-diffuion equations* 2nd Ed., Springer Verlag Berlin, Heidelberg, New-York (1993).
[25] R. A. Todor, A new approach to energy-based sparse finite-element spaces, IMA Journal of Numerical Analysis **29**(2009)(1), pp. 72-85.
[26] J. Tryoen, O. Le Maitre, M. Ndjinga and A. Ern. *Instrusive projection methods with upwinding for uncertain non-linear hyperbolic systems* Preprint, 2010.
[27] X. Wan and G. E. Karniadakis. *Long-term behaviour of polynomial chaos in stochastic flow simulations*, Comput. Meth. Appl. Mech. Engg. **195**(2006), pp. 5582-5596.
[28] W.P. Ziemer. *Weakly Differentiable Functions*, Graduate Texts in Mathematics, Springer Verlag Berlin, Heidelberg, New-York (1989).

(Siddhartha Mishra)
Seminar for Applied Mathematics
ETH
HG G. 57.2,
Rämistrasse 101, Zürich, Switzerland.
E-mail address: smishra@sam.math.ethz.ch

(Christoph Schwab)
Seminar for Applied Mathematics
ETH
HG G. 57.1,
Rämistrasse 101, Zürich, Switzerland.
E-mail address: smishra@sam.math.ethz.ch

Research Reports

No.	Authors/Title
10-24	*S. Mishra and Ch. Schwab* Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random intitial data
10-23	*J. Li, J. Xie and J. Zou* An adaptive finite element method for distributed heat flux reconstruction
10-22	*D. Kressner* Bivariate matrix functions
10-21	*C. Jerez-Hanckes and J.-C. Nédélec* Variational forms for the inverses of integral logarithmic operators over an interval
10-20	*R. Andreev* Space-time wavelet FEM for parabolic equations
10-19	*V.H. Hoang and C. Schwab* Regularity and generalized polynomial chaos approximation of parametric and random 2nd order hyperbolic partial differential equations
10-18	*A. Barth, C. Schwab and N. Zollinger* Multi-Level Monte Carlo Finite Element method for elliptic PDE's with stochastic coefficients
10-17	*B. Kågström, L. Karlsson and D. Kressner* Computing codimensions and generic canonical forms for generalized matrix products
10-16	*D. Kressner and C. Tobler* Low-Rank tensor Krylov subspace methods for parametrized linear systems
10-15	*C.J. Gittelson* Representation of Gaussian fields in series with independent coefficients
10-14	*R. Hiptmair, J. Li and J. Zou* Convergence analysis of Finite Element Methods for $H(\mathrm{div};\Omega)$-elliptic interface problems
10-13	*M.H. Gutknecht and J.-P.M. Zemke* Eigenvalue computations based on IDR
10-12	*H. Brandsmeier, K. Schmidt and Ch. Schwab* A multiscale hp-FEM for 2D photonic crystal band