ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for
Applied
Mathematics

SAM

Multi-Level Monte Carlo Finite Volume methods for uncertainty quantification of acoustic wave propagation in random heterogeneous layered medium

S. Mishra and Ch. Schwab and J. Sukys

MULTI-LEVEL MONTE CARLO FINITE VOLUME METHODS FOR UNCERTAINTY QUANTIFICATION OF ACOUSTIC WAVE PROPAGATION IN RANDOM HETEROGENEOUS LAYERED MEDIUM

S. MISHRA, CH. SCHWAB, AND J. ŠUKYS

ABSTRACT. We consider the very challenging problem of efficient uncertainty quantification for acoustic wave propagation in a highly heterogeneous, possibly layered, random medium, characterized by possibly anisotropic, piecewise log-exponentially distributed Gaussian random fields. A multi-level Monte Carlo finite volume method is proposed, along with a novel, bias-free upscaling technique that allows to represent the input random fields, generated using spectral FFT methods, efficiently. Combined together with a novel, dynamic load balancing algorithm that scales to massively parallel computing architectures, the proposed method is able to robustly compute uncertainty for highly realistic random subsurface formations that can contain a very high number (millions) of sources of uncertainty. Numerical experiments, in both two and three space dimensions, illustrating the efficiency of the method are presented.

1. INTRODUCTION

1.1. The model. The study of propagation of acoustic waves in heterogeneous media is of crucial importance to geophysics, particularly in seismic imaging. The motion of seismic waves is modeled by the linear wave equation, reading as

$$(1.1) \qquad p_{tt}(\mathbf{x}, t) - \operatorname{div}(c(\mathbf{x})\nabla p(\mathbf{x}, t)) = 0, \quad \mathbf{x} \in \mathbf{D}, \ t > 0 .$$

Here, p denotes the acoustic pressure variable and the material coefficient $c : \mathbf{D} \mapsto \mathbb{R}_+$ describes the (positive) speed of sound in a heterogeneous medium at a given domain point $\mathbf{x} \in \mathbf{D}$. Throughout this paper we consider domain $\mathbf{D}$ to be either $\mathbf{D} = \mathbb{R}^d$ (a "Cauchy problem") or a bounded axiparallel box $\mathbf{D} = I_1 \times \cdots \times I_d \subset \mathbb{R}^d$.

We rewrite the Cauchy problem for the wave equation (1.1) as a first-order linear, hyperbolic system of the form

$$(1.2) \qquad \begin{cases} p_t(\mathbf{x}, t) - \operatorname{div}(c(\mathbf{x})\mathbf{u}(\mathbf{x}, t)) = 0, \\ \quad\quad\quad \mathbf{u}_t(\mathbf{x}) - \nabla p(\mathbf{x}) = 0, \end{cases} \quad \mathbf{x} \in \mathbf{D}, \ t > 0.$$

1

This first-order system of PDEs is the prototypical example for general linear hyperbolic systems of conservation (balance) laws given by

$$
(1.3) \qquad \begin{cases} \mathbf{U}_t(\mathbf{x}, t) + \displaystyle\sum_{r=1}^{d} \frac{\partial}{\partial \mathbf{x}_r} \Big(\mathbf{A}_r(\mathbf{x}) \mathbf{U}(\mathbf{x}, t) \Big) = \mathbf{S}(\mathbf{x}, t), \\ \qquad\qquad\qquad\qquad\qquad \mathbf{U}(\mathbf{x}, 0) = \mathbf{U}_0(\mathbf{x}), \end{cases} \quad \forall (\mathbf{x}, t) \in \mathbf{D} \times \mathbb{R}_+ .
$$

Here, $\mathbf{U} : \mathbf{D} \times \mathbb{R}_+ \to \mathbb{R}^m$ denotes the vector of conserved variables, $\mathbf{A}_r : \mathbb{R}^m \to \mathbb{R}^m$ denote linear maps (fluxes), and $\mathbf{S} : \mathbf{D} \times \mathbb{R}_+ \to \mathbb{R}^m$ denotes the source term. If the source term $\mathbf{S} = 0$ (as in (1.2)), the balance law (1.3) is termed a *conservation law*.

We observe that the wave equation (1.2) is of the form (1.3) with $\mathbf{U} = (p, \mathbf{u})$, $\mathbf{S} \equiv 0$ and entries $(\mathbf{A}_r(\mathbf{x}))_{i,j}$ of matrices $\mathbf{A}_r$ for $i, j = 1, \ldots, m$ given by

$$
(1.4) \qquad (\mathbf{A}_r(\mathbf{x}))_{i,j} = \begin{cases} -c(\mathbf{x}) & i = 1, \; j = r + 1, \\ -1 & i = r + 1, \; j = 1, \\ 0 & \text{else}. \end{cases}
$$

The methodology of the present paper also extends to systems (1.3) in general, polyhedral domains with suitable boundary conditions (which, in the case of (1.1), allow first order system reformulation (1.2).

Other examples for the linear systems of conservation laws (1.3) are the equations of linear elasticity and linearized Shallow water equations of oceanography.

Given the lack of explicit solution formulas (particularly for variable coefficients and in several space dimensions), numerical methods are widely used to approximate (1.3) and, in particular, the wave equation (1.2). Popular discretization methods include finite difference, finite volume and discontinuous Galerkin methods, see [22, 54, 17] and references therein.

1.2. Uncertainty quantification.
These numerical methods require the specification of the coefficient matrices, initial data, source terms and boundary data as inputs. However, these quantities are often determined by measurements, which are typically uncertain and provide only *statistical* information about the input data. For instance, in the propagation of acoustic waves in the subsurface, the wave speeds $c(x)$ in (1.2) depend on the material properties (rock permeability) of the subsurface medium. The relative scarcity of seismic measurements leads to statistical descriptions (using available a-priori knowledge) of material properties of the medium, which then result in uncertain wave propagation speed $c(x)$. The efficient computation of the resulting solution uncertainty, given the statistical description of input uncertainty, is the central theme of uncertainty quantification (UQ).

A necessary prerequisite in UQ is an appropriate mathematical notion of random solutions for linear hyperbolic systems. In [47], we recently provided a mathematical framework and proofs of existence and uniqueness of these random solutions, including the quantification of their (spatio-temporal and statistical) regularity. There has been tremendous amount of recent interest in devising efficient UQ methods for PDEs. Among the most popular methods (particularly for elliptic and parabolic PDEs) are the stochastic Galerkin methods based on generalized Polynomial Chaos (gPC for short). An incomplete list of references on gPC methods for uncertainty quantification in hyperbolic conservation laws includes

[2, 5, 23, 24, 49, 50, 40, 52, 15] and other references therein. Although these *deterministic* methods show some promise, they suffer from the disadvantage that they are highly *intrusive*: existing codes for computing deterministic solutions of balance (conservation) laws need to be completely reconfigured for implementation of the gPC based stochastic Galerkin methods. An alternative class of methods for quantifying uncertainty in PDEs are the stochastic collocation methods, see [55, 25, 53]. In particular, a stochastic collocation method for the wave equation, even with discontinuous random speed, was recently proposed in [32]. Stochastic collocation methods are non-intrusive and easier to parallelize than the gPC based stochastic Galerkin methods. However, the lack of regularity of the solution with respect to the stochastic variables (the solution can be discontinuous in the stochastic variables if the inputs are discontinuous) impedes efficient performance of both the stochastic Galerkin as well as the stochastic collocation methods. A variant is the recently developed stochastic Finite Volume Method [26] which can deal with low parametric regularity of the computed solution, but is generally not efficient when there are a large number of sources of uncertainty (resp. a high-dimensional parameter space). Yet another set of alternative methods, which heavily rely on the assumed low "effective" number of stochastic dimensions, include adaptive analysis of variance (ANOVA) [56], proper generalized decomposition (PGD)[6] and Fokker-Planck-Kolmogorov type [51] techniques. In addition to the assumption that the "effective" number of sources of uncertainty is low, these methods require very complex representations of the input random fields, which are in practice are rarely available and for which the numerical estimation can be computationally expensive.

The afore described methods generally fail to efficiently approximate problems that either possess low stochastic regularity or a large number of sources of uncertainty or both. In fact, designing efficient methods for UQ in the wave equation with heterogenous uncertain wave speeds is highly challenging as realistic statistical representations of the underlying heterogenous media, require the use of a very large number of sources of uncertainty (stochastic dimensions) and result in wave speed fields with rather low stochastic regularity, which propagates into the solution.

For such problems, with very low parametric (stochastic) regularity, a class of viable non-intrusive methods are *Monte Carlo* (MC) methods. There, the underlying deterministic PDE is solved repeatedly for each statistical *sample* and the samples are combined to ascertain statistical information. However, robustness of MC methods with respect to solution regularity comes at the price of a low (and non-improvable) error convergence rate of $1/2$ with respect to the number of samples: a large number of numerical solves of (1.3) is required. Slow convergence has inspired the development of *Multi-Level Monte Carlo* (MLMC) methods. They were introduced by S. Heinrich for numerical quadrature [19], developed by M. Giles for Itô SPDE [11], and applied to various SPDEs [3, 8, 12, 34]. In particular, recent papers [27, 29, 28] extended the MLMC algorithm to nonlinear conservation laws. Massively parallel simulations of the random multi-dimensional Euler, magnetohydrodynamics (MHD) and shallow water equations were performed using novel static and adaptive load balancing techniques [46, 45].

1.3. **Aims and scope of the current paper.** The concept of random solutions for linear hyperbolic systems (1.3) was developed and shown to be well-posed in a recent paper [47]. Furthermore, we also outlined the basic framework of MC and

MLMC methods for computational uncertainty quantification for (1.3). However, the numerical examples presented in [47] were rather elementary. Major algorithmic adaptations are required in order to apply the MLMC method on large scale two and three dimensional problems with realistic representations of the underlying uncertainty in the properties of the subsurface. The main aim of the current paper is to propose these adaptations and extend the MLMC method to quantity uncertainty in such realistic configurations. To this end,

- we consider a realistic statistical description of the random material coefficients $c(x)$, given by a piecewise log-gaussian, anisotropic random field. Although MLMC methods for *stationary* problems with a approximate representation (via truncated Karhunen-Loève expansions) of *isotropic, loggaussian* rough coefficients were considered in [48, 21], a key aspect of the current paper deals with the dynamic problem of wave propagation in random media when anisotropically correlated, layered log-gaussian statistical models of the subsurface medium are assumed.
- Although independent identically distributed (i.i.d.) samples for the input layered log-gaussian permeability field can be efficiently generated using spectral FFT (Fast Fourier Transform) generator [7, 34, 35, 42], one needs to suitably couple this approximate representation of a complex random medium with the MLMC framework in order to keep the computational cost reasonable. One principal tool developed in the current paper is a *bias-free* upscaling technique which allows to *efficiently* generate identically distributed realizations of random inputs on several mesh resolutions without accessing finer resolutions.
- The explicit timestepping used in the pathwise simulations entails, via the CFL condition, a *sample-dependent timestep* which, in turn, implies delicate load-balancing and scaling issues within the MLMC algorithm, particularly for a very large number of processors. We couple a novel dynamic load balancing algorithm designed recently by one of the authors in [45] with the MLMC framework.

Thus, we combine the MLMC method with a novel bias-free, upscaled representation of input random fields and an adaptive load balancing procedure in the current paper to obtain a non-intrusive and highly effective computational UQ framework for the propagation of acoustic waves in realistic, highly heterogenous, possibly discontinuous, two- and three-dimensional random media.

The rest of the paper is organized as follows. In section 2, we recapitulate the main theoretical results from [47]: we present the stochastic linear system of conservation laws, define random weak solutions and state the results on their existence and uniqueness. In section 3, we present the MLMC-FVM algorithm and provide the asymptotic bounds for error vs. computational work. In section 4, the spectral FFT generator for log-gaussian, anisotropically correlated layered material coefficients is presented, together with the bias-free upscaling technique required for the coupling of the spectral generator to MLMC framework. In section 5, parallel implementation ALSVID-UQ of MLMC-FVM is discussed. In section 6, the acoustic wave equation (1.2) will be considered in two and three spatial dimensions, with random log-normally distributed highly heterogeneous layered material coefficients $c(\mathbf{x})$.

2. Systems of stochastic linear hyperbolic conservation laws

We fix notation and describe the probabilistic modeling of random inputs and solution by introducing random fields. Next, several well-posedness results from [47] are recapitulated and applied to a special case of the acoustic wave equation (1.2).

2.1. **Preliminaries.** For the mathematical description of uncertainty in inputs and solutions of PDEs, as well as for stating convergence results of MC methods, we place ourselves into the setting of Kolmogorov's probability theory, and recapitulate basic terminlogy and notation which is used throughout this paper, from [41]. Uncertain inputs for (1.3) then take the form of random fields with prescribed probability laws.

Let $(\Omega, \mathcal{F})$ be a measurable space, with Ω denoting the set of all elementary events $\omega \in \Omega$, and $\mathcal{F}$ a σ-algebra of all possible events in our probability model. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a 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. Denoting a second measurable space by $(E, \mathcal{G})$, an *E-valued random field* (or random variable taking values in E) is any mapping $X : \Omega \to E$ such that the set $\{\omega \in \Omega : X(\omega) \in A\} = \{X \in A\} \in \mathcal{F}$ for any $A \in \mathcal{G}$, i.e. such that X is a $\mathcal{G}$-measurable mapping $\Omega \to E$. $\mathcal{L}(X)$ denotes the *law of X under* $\mathbb{P}$,

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

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

Definition 2.1 (Random field). *For a separable Banach space E, an E-valued random field is a $(\mathcal{F}, \mathcal{B}(E))$-measurable mapping $\Omega \to E$,*

$$X : (\Omega, \mathcal{F}) \to (E, \ \mathcal{B}(E)).$$

Lemma 2.2 (Random field norm is a random variable [41])**.** *Let E be a separable Banach space and let $X : \Omega \to E$ be an E-valued random field on $(\Omega, \mathcal{F})$. Then the mapping $\Omega \ni \omega \mapsto \|X(\Omega)\|_E \in \mathbb{R}$ is measurable, i.e. it is a random variable $(\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

As E is separable, the random field $X : \Omega \to E$ is *Bochner integrable* with respect to the probability measure $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{F})$,

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

By $L^1(\Omega, E) = L^1((\Omega, \mathcal{F}, \mathbb{P}), E)$ we denote the set of all (equivalence classes of) Bochner integrable, E-valued random fields X, equipped with the norm

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

More generally, for $1 \le p < \infty$, we define $L^p(\Omega, E) = L^p((\Omega, \mathcal{F}, \mathbb{P}), E)$ as the set of Bochner p-integrable random fields taking values E, equipped with the norm

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

For $p = \infty$, we denote by $L^\infty(\Omega, E) = L^\infty((\Omega, \mathcal{F}, \mathbb{P}), E)$ the set of all E-valued random fields which are $\mathbb{P}$-almost surely bounded, equipped with the norm

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

For $X \in L^1(\Omega, X)$, the mathematical *expectation* (or "ensemble average") of an E-valued random field X is well-defined and denoted by $\mathbb{E}[X]$.

2.2. **Random conservation laws.** As the inputs to (1.3) are random fields, so are the solutions of the balance law (1.3). Statistical moments of the random field solution such as the expectation $\mathbb{E}[\mathbf{U}]$ and spatiotemporal correlations (covariance functions) are typical (deterministic!) quantities of interest. A realization $\omega \in \Omega$ of the linear system of balance laws (1.3) with random initial data, coefficients and sources is given by

$$(2.1) \quad \begin{cases} \mathbf{U}_t(\mathbf{x}, t, \omega) + \displaystyle\sum_{r=1}^{d} \frac{\partial}{\partial \mathbf{x}_r}\Big(\mathbf{A}_r(\mathbf{x}, \omega)\mathbf{U}\Big) = \mathbf{S}(\mathbf{x}, t, \omega), \\ \qquad\qquad\qquad\qquad \mathbf{U}(\mathbf{x}, 0, \omega) = \mathbf{U}_0(\mathbf{x}, \omega), \end{cases} \quad \forall(\mathbf{x}, t) \in \mathbf{D} \times \mathbb{R}_+ .$$

Here, the initial data $\mathbf{U}_0$ is a $\mathbf{L}^2(\mathbf{D})$-valued random field

$$\mathbf{U}_0 : (\Omega, \mathcal{F}) \to \big(\mathbf{L}^2(\mathbf{D}), \mathcal{B}(\mathbf{L}^2(\mathbf{D}))\big), \quad \Omega \ni \omega \mapsto \mathbf{U}_0(\mathbf{x}, \omega),$$

the matrices $\mathbf{A}_r$ are $(C^1(\mathbf{D}))^{m \times m}$-valued random fields,

$$\mathbf{A}_r : (\Omega, \mathcal{F}) \to (C^1(\mathbf{D})^{m \times m}, \mathcal{B}(C^1(\mathbf{D})^{m \times m})), \quad \Omega \ni \omega \mapsto \mathbf{A}_r(\mathbf{x}, \omega),$$

the source $\mathbf{S}$ is a $C_b\big([0, T], \mathbf{L}^2(\mathbf{D} \times \mathbb{R}^m)\big)$-valued random field,

$$\mathbf{S} : (\Omega, \mathcal{F}) \to \big(C_b\big([0, T], \mathbf{L}^2(\mathbf{D} \times \mathbb{R}^m)\big), \mathcal{B}\big(C_b\big([0, T], \mathbf{L}^2(\mathbf{D} \times \mathbb{R}^m)\big)\big)\big),$$

and the solution $\mathbf{U}$ is a $C_b([0, T], \mathbf{L}^2(\mathbf{D}))$-valued random field,

$$\mathbf{U} : (\Omega, \mathcal{F}) \to \big(C_b\big([0, T], \mathbf{L}^2(\mathbf{D})\big), \ \mathcal{B}\big(C_b\big([0, T], \mathbf{L}^2(\mathbf{D})\big)\big)\big), \quad \Omega \ni \omega \mapsto \mathbf{U}(\mathbf{x}, t, \omega).$$

In the case of the wave equation (1.2), randomness in the coefficient matrices $\mathbf{A}_r$ is inherited (through relation (1.4)) from the underlying random material coefficients $c(\mathbf{x}, \omega)$, modeled as a $C^1(\mathbf{D})$-valued random field,

$$\mathbf{c} : (\Omega, \mathcal{F}) \to \big(C^1(\mathbf{D}), \mathcal{B}(C^1(\mathbf{D}))\big), \quad \Omega \ni \omega \mapsto c(\mathbf{x}, \omega).$$

For simplicity of exposition, all definitions and theoretical results will be stated explicitly only for Cauchy problems with $\mathbf{D} = \mathbb{R}^d$. Analogous results are available for general bounded domains, for instance, for periodic bounded Cartesian domains $\mathbf{D} = I_1 \times \cdots \times I_d$.

Following [47], we define the notion of solutions of (2.1), which is a stochastic extension of conventional definition from [22, 17] of weak solutions for deterministic case in (1.3).

Definition 2.3 (Random weak solution)**.** *A random field* $\mathbf{U} : \Omega \ni \omega \mapsto \mathbf{U}(\mathbf{x}, t, \omega)$, *i.e. a measurable mapping from* $(\Omega, \mathcal{F})$ *to* $C([0, T], \mathbf{L}^2_{\mathrm{loc}}(\mathbb{R}^d))$, *is random weak solution to the stochastic conservation law* (2.1) *on* $\mathbf{D} = \mathbb{R}^d$ *if it is* $\mathbb{P}$-a.s. *a weak*

solution of (1.3) *on* $\mathbf{D} = \mathbb{R}^d$, *i.e.* $\mathbf{U}$ *satisfies the variational formulation for all test functions* $\boldsymbol{\varphi} \in \mathbf{C}_c^1(\mathbb{R}^d \times \mathbb{R}_+)$ *with notation* $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$ *and* $\mathbb{P}$-*a.e.* $\omega \in \Omega$,

(2.2)
$$\int_{\mathbb{R}^d \times \mathbb{R}_+} \left(\mathbf{U} \cdot \boldsymbol{\varphi}_t + \sum_{r=1}^d \mathbf{A}_r \mathbf{U} \cdot \frac{\partial}{\partial \mathbf{x}_i} \boldsymbol{\varphi} \right) d\mathbf{x} \, dt + \int_{\mathbb{R}^d} \mathbf{U}_0 \cdot \boldsymbol{\varphi}|_{t=0} \, d\mathbf{x} = \int_{\mathbb{R}^d \times \mathbb{R}_+} \mathbf{S} \cdot \boldsymbol{\varphi} \, d\mathbf{x} \, dt.$$

Next, we specify the notion of hyperbolicity for a general linear system of random conservation laws (1.3).

Definition 2.4 (Hyperbolicity)**.** *Denote a convex combination of random matrices* $\mathbf{A}_r(\mathbf{x}, \omega)$ *by*

$$\mathbf{P}^{\mathbf{w}} : \mathbb{R}^m \times \Omega \to \mathbb{R}^m \times \Omega, \quad \mathbf{P}^{\mathbf{w}}(\mathbf{x}, \omega) = \sum_{r=1}^d \mathbf{w}_r \mathbf{A}_r(\mathbf{x}, \omega), \quad \mathbf{w} \in \mathbb{S}^{d-1}.$$

Consider the eigen-decompositions of all possible $\mathbf{P}^{\mathbf{w}}(\mathbf{x}, \omega)$, *i.e. for all* $\mathbf{w} \in \mathbb{S}^{d-1}$,

(2.3)
$$\mathbf{P}^{\mathbf{w}}(\mathbf{x}, \omega) = \mathbf{Q}^{\mathbf{w}}(\mathbf{x}, \omega) \mathbf{\Lambda}^{\mathbf{w}}(\mathbf{x}, \omega) \mathbf{Q}^{\mathbf{w}}(\mathbf{x}, \omega)^{-1}, \quad \mathbf{\Lambda}^{\mathbf{w}}(\mathbf{x}, \omega) = \mathrm{diag}(\lambda_1^{\mathbf{w}}, \ldots, \lambda_m^{\mathbf{w}})(\mathbf{x}, \omega),$$

where $\mathbf{\Lambda}^{\mathbf{w}}(\mathbf{x}, \omega)$ *are the diagonal matrices containing the eigenvalues* $\lambda_1^{\mathbf{w}}, \ldots, \lambda_m^{\mathbf{w}}$, *and* $\mathbf{Q}^{\mathbf{w}}(\mathbf{x}, \omega)$ *are the corresponding similarity transformation matrices containing eigenvectors as columns. The random linear system of conservation laws* (2.1) *is* $\mathbb{P}$-*a.s. hyperbolic* [47] *if for all* $\mathbf{x} \in \mathbb{R}^d$ *all eigenvalues* $\lambda_1^{\mathbf{w}}(\mathbf{x}, \omega), \ldots, \lambda_m^{\mathbf{w}}(\mathbf{x}, \omega)$ *are real* $\mathbb{P}$-*a.s. and there exists* $K(\omega) < \infty$ *such that*

(2.4)
$$\sup_{\mathbf{x} \in \mathbf{D}, \ \mathbf{w} \in \mathbb{S}^{d-1}} \|\mathbf{Q}^{\mathbf{w}}(\mathbf{x}, \omega)^{-1}\| \|\mathbf{Q}^{\mathbf{w}}(\mathbf{x}, \omega)\| \leq K(\omega), \quad \forall \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

Next, we present the result from [47] on the well-posedness of (2.1).

Theorem 2.5. *In the random linear system of balance laws* (2.1) *on* $\mathbf{D} = \mathbb{R}^d$, *assume that the following holds for some* $k \in \mathbb{N} \cup \{0, \infty\}$:

(A1) (2.1) *is hyperbolic with constant* $K(\omega) < \infty$ *in* (2.4) *for* $\mathbb{P}$-*a.e.* $\omega \in \Omega$, *and*

(2.5)
$$\bar{K}_k = \|K(\omega)\|_{L^k(\Omega, \mathbb{R})} < \infty,$$

(A2) there exist $r_{\mathbf{0}}, r_{\mathbf{S}} \in \mathbb{N} \cup \{0, \infty\}$ *and* $r_{\mathbf{A}} \in \mathbb{N} \cup \{\infty\}$ *such that for both* $p \in \{2, \infty\}$,

(2.6)
$$\mathbf{U}_0 \in L^k(\Omega, \mathbf{W}^{r_{\mathbf{0}}, p}(\mathbb{R}^d)),$$
$$\mathbf{S} \in L^k(\Omega, \mathbf{W}^{r_{\mathbf{S}}, p}(\mathbb{R}^d)),$$
$$\mathbf{A}_r \in L^0(\Omega, (C^{r_{\mathbf{A}}}(\mathbb{R}^d))^{m \times m}),$$

(A3) each random field $\mathbf{A}_r$, $r = 1, \ldots, d$, *is independent of* $\mathbf{U}_0$ *and* $\mathbf{S}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$. *Then, for* $T < \infty$, (2.1) *admits a unique random weak solution*

(2.7)
$$\mathbf{U} : \Omega \to C([0, T], \mathbf{L}^2(\mathbb{R}^d)), \qquad \omega \mapsto \mathbf{U}(\cdot, \cdot, \omega), \quad \forall \omega \in \Omega,$$

where $\mathbf{U}(\cdot, \cdot, \omega)$ *is the solution to the* deterministic *system* (1.3)*. Moreover,* $\forall t \in [0, T]$,

(2.8)
$$\|\mathbf{U}(\cdot, t, \omega)\|_{\mathbf{L}^2(\mathbb{R}^d)} \leq \|K(\omega), \mathbf{U}_0(\cdot, \omega), \ \mathbf{S}(\cdot, \omega), \ t\|_{\mathbf{L}^2(\mathbb{R}^d)}, \quad \mathbb{P}\text{-}a.s.,$$

(2.9)
$$\|\mathbf{U}\|_{L^k(\Omega,C([0,T],\mathbf{L}^2(\mathbb{R}^d)))} \le \|K,\mathbf{U}_0,\mathbf{S},t\|_{L^k(\Omega,\mathbf{L}^2(\mathbb{R}^d))} = \bar{K}_k\|1,\mathbf{U}_0,\mathbf{S},t\|_{L^k(\Omega,\mathbf{L}^2(\mathbb{R}^d))},$$

with the notations (for arbitrary Banach space V)

(2.10)
$$\|K,\mathbf{U},\mathbf{S},t\|_{L^k(\Omega,V)} := \|\|K,\mathbf{U},\mathbf{S},t\|_V\|_{L^k(\Omega,\mathbb{R})}, \quad \|K,\mathbf{U},\mathbf{S},t\|_V := K(\|\mathbf{U}\|_V + |t|\|\mathbf{S}\|_V).$$

Furthermore, the pathwise regularity of $\mathbf{U}$ is given by $\bar{r} = \min\{r_\mathbf{0}, r_\mathbf{S}, r_\mathbf{A}\}$, leading to

(2.11)
$$\mathbf{U} \in L^k(\Omega, C([0,T], \mathbf{W}^{\bar{r},p}(\mathbb{R}^d))), \quad p \in \{2, \infty\}.$$

For the proof we refer to [47]. This theorem ensures the existence of the statistical k-th moments [27] of the random weak solution $\mathbf{U}$, provided $\mathbf{U}_0, \mathbf{S} \in L^k(\Omega, \mathbf{L}^2(\mathbb{R}^d))$ and $K \in L^k(\Omega)$.

2.3. **Hyperbolicity of the random acoustic wave equation.** Next, we briefly investigate the hyperbolicity of the wave equation (1.1) as a random linear system of conservation laws (2.1) in order to ensure that the hypothesis (2.5) in Theorem 2.5 holds. Furthermore, we derive a condition on the material coefficient $\mathbf{c}$ in order to attain finite expected maximum wave propagation speeds, i.e. which ensure that (3.11) holds. We analyze the structure of the matrices $\mathbf{A}_r$ in (2.1), which define a strongly hyperbolic linear system of conservation laws under some conditions on the statistical properties of the material coefficient $\mathbf{c}$. For brevity of exposition, we consider the one-dimensional case $d = 1$. Then, for each $\omega \in \Omega$ and $\mathbf{x} \in \mathbf{D}$, there exists an invertible matrix $\mathbf{Q}(\mathbf{x}, \omega)$,

$$\mathbf{Q}(\mathbf{x}, \omega) = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{c(\mathbf{x},\omega)}} & -\frac{1}{\sqrt{c(\mathbf{x},\omega)}} \\ 1 & 1 \end{bmatrix},$$

diagonalizing the matrix $\mathbf{A}(\mathbf{x}, \omega)$ defined in (1.4),

$$\mathbf{Q}(\mathbf{x}, \omega)\mathbf{A}_1(\mathbf{x}, \omega)\mathbf{Q}(\mathbf{x}, \omega)^{-1} = \begin{bmatrix} -\sqrt{c(\mathbf{x}, \omega)} & 0 \\ 0 & \sqrt{c(\mathbf{x}, \omega)} \end{bmatrix}.$$

The maximum wave speed $\lambda(\omega)$ is then given by

(2.12)
$$\lambda(\omega) = \|\sqrt{c(\cdot, \omega)}\|_{C(\mathbf{D})}.$$

Since

$$\|\mathbf{Q}_\mathbf{x}(\omega)\|\|\mathbf{Q}_\mathbf{x}^{-1}(\omega)\| = \max\{c^{\frac{1}{2}}, c^{-\frac{1}{2}}\} \le c^{\frac{1}{2}} + c^{-\frac{1}{2}},$$

the uniform boundedness

(2.13)
$$c, c^{-1} \in L^\infty(\Omega, C(\mathbf{D}))$$

ensures $\bar{K}_\infty < \infty$. However, assumption (2.13) is overly strict and can be further relaxed. In particular, for $1 \le k < \infty$,

(2.14)
$$c^{\frac{1}{2}}, c^{-\frac{1}{2}} \in L^k(\Omega, C(\mathbf{D}))$$

implies

$$K^k(\omega) = \|\max\{c^{\frac{1}{2}}(\cdot, \omega), c^{-\frac{1}{2}}(\cdot, \omega)\}\|_{C(\mathbf{D})}^k \le \|c(\cdot, \omega)\|_{C(\mathbf{D})}^{\frac{k}{2}} + \|c^{-1}(\cdot, \omega)\|_{C(\mathbf{D})}^{\frac{k}{2}}.$$

Hence $c^{\frac{1}{2}}, c^{-\frac{1}{2}} \in L^k(\Omega, C(\mathbf{D}))$ is sufficient to ensure (A1) of Theorem 2.5:

(2.15)
$$\bar{K}_k^k = \mathbb{E}[K^k(\omega)] \le \|c(\cdot, \omega)^{\frac{1}{2}}\|_{L^k(\Omega,C(\mathbf{D}))}^k + \|c^{-\frac{1}{2}}(\cdot, \omega)\|_{L^k(\Omega,C(\mathbf{D}))}^k < \infty.$$

Since the non-zero eigenvalues of $\mathbf{A}_r \in \mathbb{R}^{m \times m}$ are $\pm\sqrt{c(\mathbf{x}, \omega)}$, the assumption of finite expected maximum wave speed $\bar{\lambda}$ in (3.11) of the forthcoming subsection 3.1 holds, provided $\sqrt{c} \in L^1(\Omega, C(\mathbf{D}))$,

$$(2.16) \qquad \bar{\lambda} := \mathbb{E}[\lambda] = \|\sqrt{c}\|_{L^1(\Omega, C(\mathbf{D}))} < \infty, \quad \text{for} \quad \sqrt{c} \in L^1(\Omega, C(\mathbf{D})).$$

Finally, assumption (A2) in the Theorem 2.5 holds with $r_{\mathbf{A}} = r_c$, provided

$$c \in L^0(\Omega, C^{r_c}(\mathbf{D})),$$

since all non-constant entries in matrices $\mathbf{A}_r$ are simply given by the (negated) coefficient $\mathbf{c}(\mathbf{x}, \omega)$, see (1.4).

3. Multi-Level Monte Carlo Finite Volume method

Next, we aim to design an efficient numerical scheme to approximate solution statistics for the stochastic balance law (2.1). This entails discretizing spatio-temporal space $\mathbf{D} \times [0, T]$ (for bounded domains $\mathbf{D}$ and finite time horizon $T < \infty$) as well as the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the first subsection, we begin with the Finite Volume Method (FVM) for the spatio-temporal space. Afterwards, we present Monte Carlo (MC) and a more efficient Multi-Level Monte Carlo (MLMC) sampling type algorithms for the discretization of the probability space and provide convergence results.

3.1. **Finite Difference and Finite Volume methods.** For complicated coefficient matrices $\mathbf{A}_r$ (or material coefficients c in the case of the acoustic wave equation), exact analytic solutions to deterministic systems of linear balance laws (1.3) are not available. For continuous $\mathbf{U}_0$ and $\mathbf{S}$ (then the solution $\mathbf{U}$ is also continuous), conventional Finite Difference methods [17, 54] can be used where the numerical scheme is obtained by approximating spatial and temporal derivatives in (1.3) using upwinded difference quotients. For discontinuous $\mathbf{U}_0$ and $\mathbf{S}$, (then the solution $\mathbf{U}$ is also discontinuous) we present the Finite Volume Method [22].

Consider a bounded axiparallel domain $\mathbf{D} = I_1 \times \cdots \times I_d \subset \mathbb{R}^d$ and let $\mathcal{T} = \mathcal{T}^1 \times \cdots \times \mathcal{T}^d$ denote a uniform axiparallel quadrilateral mesh of the domain $\mathbf{D}$, consisting of identical cells $C_{\mathbf{j}} = C_{\mathbf{j}_1} \times \cdots \times C_{\mathbf{j}_d}$, $\mathbf{j}_r = 1, \ldots, \#\mathcal{T}^r$.

Assume *mesh widths* are equal in each dimension, i.e. $\Delta x := \frac{|I_1|}{\#\mathcal{T}_1} = \cdots = \frac{|I_d|}{\#\mathcal{T}_d}$. Define the approximations to cell averages of the solution $\mathbf{U}$ and source term $\mathbf{S}$ by

$$\mathbf{U}_{\mathbf{j}}(t) \approx \frac{1}{|C_{\mathbf{j}}|} \int_{C_{\mathbf{j}}} \mathbf{U}(\mathbf{x}, t) d\mathbf{x}, \quad \mathbf{S}_{\mathbf{j}} \approx \frac{1}{|C_{\mathbf{j}}|} \int_{C_{\mathbf{j}}} \mathbf{S}(\mathbf{x}) d\mathbf{x}.$$

Then, a semi-discrete finite volume scheme [22] for approximating (1.3) is given by

$$(3.1) \qquad \partial_t \mathbf{U}_{\mathbf{j}}(t) = -\sum_{r=1}^{d} \frac{1}{\Delta x} \left(\mathbf{F}_{\mathbf{j} + \frac{1}{2}\mathbf{e}_r} - \mathbf{F}_{\mathbf{j} - \frac{1}{2}\mathbf{e}_r} \right) - \mathbf{S}_{\mathbf{j}},$$

where *numerical fluxes* $\mathbf{F}^r$ are defined by using (approximate) solutions of local Riemann problems (in direction r) at each cell interface. High order accuracy is achieved by using non-oscillatory TVD, ENO, WENO methods [16, 18]. At time steps t^n, approximations $\mathbf{U}_{\mathcal{T}}^n = \mathbf{U}_{\mathcal{T}}(\mathbf{x}, t^n) = \mathbf{U}_{\mathbf{j}}(t^n)$ for $\mathbf{x} \in C_{\mathbf{j}}$ are obtained by SSP Runge-Kutta methods, where the time step size Δt is limited by the CFL condition (with λ_m being eigenvalues from (2.3)),

$$(3.2) \qquad \lambda_{\max} \frac{\Delta t}{\Delta x} \leq \frac{1}{2}, \quad \lambda_{\max} := \max_{\mathbf{j}} \max_{m} \left| \lambda_m(\mathbf{x}_{\mathbf{j}}, \omega) \right|.$$

Fluxes on the boundary of the computational domain are computed using so-called ghost cells [22], denoted (in the direction r) by $\mathcal{C}_{\mathbf{j}+\mathbf{e}_r}$ for $\mathbf{j}_r = \#\mathcal{T}_r$ and $\mathcal{C}_{\mathbf{j}-\mathbf{e}_r}$ for $\mathbf{j}_r = 1$, which are "outside" of the computational domain $\mathbf{D}$. Periodic boundary conditions can be implemented straight-forwardly by copying ghost cell values accordingly, i.e. we set

$$(3.3) \qquad \mathbf{U}_{\mathbf{j}+\mathbf{e}_r} = \mathbf{U}_{\mathbf{j}-(\#\mathcal{T}_r-1)\mathbf{e}_r}, \quad \mathbf{U}_{\mathbf{j}-\mathbf{e}_r} = \mathbf{U}_{\mathbf{j}+(\#\mathcal{T}_r-1)\mathbf{e}_r}.$$

In the case of the acoustic wave equation (1.2), *perfectly reflecting boundary conditions* will also be used. There the acoustic pressure p and velocities $\mathbf{u}$ are set according to

$$(3.4) \qquad p_{\mathbf{j}+\mathbf{e}_r} = p_{\mathbf{j}}, \quad \mathbf{u}_{\mathbf{j}+\mathbf{e}_r} = -\mathbf{u}_{\mathbf{j}}; \qquad p_{\mathbf{j}-\mathbf{e}_r} = p_{\mathbf{j}}, \quad \mathbf{u}_{\mathbf{j}-\mathbf{e}_r} = -\mathbf{u}_{\mathbf{j}}.$$

Ghost cell values for other remaining domain-dependent input data are set analogously.

Before we proceed with the definition of the MC-FVM scheme, we assume that an abstract FVM scheme (3.1) satisfies the following assumption.

Assumption 3.1 (FVM). *We assume, that under CFL condition* (3.2), *the approximate FVM solution* $\mathbf{U}_{\mathcal{T}}^n$ *of an abstract FVM scheme* (3.1) *converges to the unique weak solution* $\mathbf{U}$ *of the linear system of conservation laws* (1.3). *Furthermore, if boundary conditions do not introduce additional energy into the solution, FVM approximation* $\mathbf{U}_{\mathcal{T}}^n$ *satisfies the energy inequality*

$$(3.5) \qquad \|\mathbf{U}_{\mathcal{T}}(\cdot, t)\|_{\mathbf{L}^2(\mathbf{D})} \leq K\|\mathbf{U}_{\mathcal{T}}^0, \mathbf{S}_{\mathcal{T}}, t\|_{\mathbf{L}^2(\mathbf{D})},$$

and the approximation error converges (as $\Delta x \to 0$*) with rate* $s > 0$*, i.e., there exists a constant* $C_s > 0$ *which is independent of* $\Delta x = \Delta x(\mathcal{T})$ *such that, as* $\Delta x(\mathcal{T}) \to 0$*, the following holds,*

$$(3.6) \qquad \|\mathbf{U}(\cdot, t) - \mathbf{U}_{\mathcal{T}}(\cdot, t)\|_{\mathbf{L}^2(\mathbf{D})} \leq C_s \Delta x^s tK\|\mathbf{U}_0, \mathbf{S}, t\|_{\mathbf{H}^s(\mathbf{D})} .$$

Here, $\mathbf{H}^s(\mathbf{D})$ *denotes the Hilbert space* $W^{s,2}(\mathbf{D})^m$ *of* s*-times weakly differentiable (equivalence classes of) vector functions with (component-wise)* $L^2(\mathbf{D})$*-integrable weak derivatives.*

Assumption 3.1 is satisfied by many standard FVM (for small s) schemes with periodic and reflective boundary conditions, we refer to [13, 14, 17, 22, 54] and the references therein for further details. In particular, the convergence estimate (3.6) is known to hold for first-order FVM schemes by results of Kusznetsov (see, e.g. [10]) with $s = 1/2$. We also assume $s = 1$ for second order FVM schemes. In general, for q-th order (formally) accurate schemes, full convergence order $s = q$ is achieved for sufficiently smooth solutions with $\bar{r} \geq q$ in (2.6), whereas irregular solutions with shocks ($\bar{r} = 0$) converge with order $s \leq 1$, equal to only *half* of the formal order q, resulting in $s = 1/2$ or $s = 1$.

3.2. **Monte Carlo Finite Volume method.** For the discretization of the stochastic space Ω, we will employ the sampling type methods. We are interested in the computational estimation of the "mean field" or "ensemble average", i.e. of $\mathcal{M}^1(\mathbf{U}) = \mathbb{E}[\mathbf{U}]$. To this end, we use the *Monte Carlo Finite Volume method* (MC-FVM) to approximate $\mathbb{E}[\mathbf{U}]$. It is based on the straightforward idea of generating independent identically distributed (i.i.d.) samples of the random input data

$$(3.7) \qquad \mathbf{I}(\omega) = \{\mathbf{U}_0(\omega), \mathbf{S}(\omega), \mathbf{A}_1(\omega), \ldots, \mathbf{A}_d(\omega)\},$$

and then performing the FVM simulation for each such sample.

Definition 3.2 (MC-FVM Scheme). *The Monte Carlo Finite Volume Method consists of the following three main steps:*

1. **Sample:** *Given mesh $\mathcal{T}$, we draw M i.i.d. input data samples $\mathbf{I}_{\mathcal{T}}^i$ with $i = 1, 2, \ldots, M$ from the random input fields $\mathbf{I}(\omega)$, which are directly approximated by piecewise constant functions as described in subsection 3.1.*
2. **Solve:** *For each approximated realization $\mathbf{I}_{\mathcal{T}}^i$ of random input data $\mathbf{I}(\omega)$, the underlying balance law (1.3) is solved numerically by the Finite Volume Method (3.1) on mesh $\mathcal{T}$ with mesh width $\Delta x = \Delta x(\mathcal{T})$. We denote the solutions by $\mathbf{U}_{\mathcal{T}}^{i,n}(\mathbf{x}) = \mathbf{U}_{\mathcal{T}}^i(\mathbf{x}, t^n)$.*
3. **Estimate Statistics:** *We estimate the expectation $\mathbb{E}[\mathbf{U}]$ of the random solution field $\mathbf{U}(\cdot, t, \omega)$ at time $t = t^n$ with the sample mean (average) of the approximate solutions,*

$$(3.8) \qquad \mathbb{E}[\mathbf{U}(\cdot, t^n, \omega)] \approx E_M[\mathbf{U}_{\mathcal{T}}^n] = E_M[\mathbf{U}_{\mathcal{T}}^n(\cdot)] := \frac{1}{M} \sum_{i=1}^{M} \mathbf{U}_{\mathcal{T}}^{i,n}(\cdot).$$

Higher statistical moments of $\mathbf{U}$ can be approximated analogously, cp. [27].

The following result from [47] addresses the convergence of MC-FVM as $M \to \infty$ and $\Delta x \to 0$.

Theorem 3.3 (MC-FVM convergence). *Consider a linear system of conservation laws (2.1) and assume that the hypothesis of Theorem 2.5 is satisfied with $k \geq 2$, i.e. second moments of the random initial data $\mathbf{U}_0$, source $\mathbf{S}$ and K exist. Assume further that we are given a FVM such that (3.2) holds and such that Assumption 3.1 is satisfied; in particular, assume that the deterministic FVM scheme converges at rate $s > 0$ as in (3.6). Then, for time $t = t^n$, the MC estimate $E_M[\mathbf{U}_{\mathcal{T}}^n(\cdot)](\omega)$ defined in (3.8) satisfies, for every M, the error bound*

$$\begin{aligned}
(3.9) \qquad \|\mathbb{E}[\mathbf{U}(t^n)] - E_M[\mathbf{U}_{\mathcal{T}}^n](\omega)\|_{L^2(\Omega, \mathbf{L}^2(\mathbf{D}))} &\leq C_s \Delta x^s \|K, \mathbf{U}_0, \mathbf{S}, t^n\|_{L^1(\Omega, \mathbf{H}^s(\mathbf{D}))} \\
&\quad + M^{-\frac{1}{2}} \|K, \mathbf{U}_0, \mathbf{S}, t^n\|_{L^2(\Omega, \mathbf{L}^2(\mathbf{D}))}.
\end{aligned}$$

where constant C_s denotes the constant in the deterministic a-priori error bounds (3.6) which is in particular independent of M and Δx; the norms are as in (2.10).

To equilibrate statistical and spatio-temporal discretization errors in (3.9), we require the number of Monte Carlo samples to equal (asymptotically)

$$(3.10) \qquad M = \mathcal{O}(\Delta x^{-2s}).$$

Next, we are interested in the asymptotic behavior of the error (3.9) vs. the computational work of all FVM solves required in (3.8). As emphasized in [47], since coefficients $\mathbf{A}_r$ directly relate to the fastest wave speed $\lambda_{\max}$ via the CFL condition (3.2), $\lambda_{\max}$ can *strongly depend* on the particular realizations of the *random* input data $\mathbf{I}^i$, for $i = 1, \ldots, M$. For the remaining sections, we assume that *the expected maximum wave speed $\bar{\lambda} = \mathbb{E}[\lambda_{\max}(\omega)]$ is finite*, i.e. we stipulate

$$(3.11) \qquad \bar{\lambda} = \mathbb{E}[\lambda_{\max}(\omega)] < \infty.$$

Under assumption (3.11), the *expected* computational work $\mathbb{E}[\text{Work}]$ for the MC-FVM estimate (3.8) is likewise *finite*, and hence the resulting error vs. *expected*

computational work $\mathbb{E}\,[\text{Work}]$ of the MC-FVM scheme (3.8) with the $L^2(\Omega)$-type error bound (3.9) is given by

$$(3.12) \qquad \|\mathbb{E}[\mathbf{U}(t^n)] - E_M[\mathbf{U}_{\mathcal{T}}^n](\omega)\|_{L^2(\Omega, \mathbf{L}^2(\mathbf{D}))} \lesssim \mathbb{E}\,[\text{Work}]^{-\frac{s}{d+1+2s}}\,.$$

Notice, that the convergence rate $s/(d+1+2s)$ of MC-FVM scheme is considerably lower than the convergence rate $s/(d+1)$ of the deterministic FVM scheme. Hence, the MC-FVM is considerably more expensive than the standard FVM for a deterministic conservation law. As an example, a first order scheme ($s = 1/2$) leads to a convergence rate of $1/6$ for the MC-FVM as compared to a convergence rate of $1/4$ for the standard FVM for a deterministic conservation law. This slow convergence entails high computational costs for MC type methods. In particular, quantifying uncertainty with MC methods for systems of conservation laws in several space dimensions becomes very costly.

3.3. **Multi-Level Monte Carlo Finite Volume method.** Given the slow convergence of MC-FVM, the multi-level Monte Carlo finite volume method (MLMC-FVM) was proposed in [47] and in related papers such as [27, 29, 28, 30]. The key idea behind MLMC-FVM is to simultaneously draw MC samples on a hierarchy of nested grids, as originally suggested by Mishra and Schwab [27] for the Finite Volume Method. The key ingredient in the Multi-level 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. The Multi Level Monte Carlo Finite Volume algorithm consists of the following four main steps:

1. **Hierarchy of space-time discretizations:** Assume that $\{\mathcal{T}_\ell\}_{\ell=0}^L$ is a family of *nested triangulations* of bounded Cartesian domain $\mathbf{D}$,

$$\{\mathcal{T}_\ell\}_{\ell=0}^L = \{\mathcal{T}_\ell^1 \times \cdots \times \mathcal{T}_\ell^d\}_{\ell=0}^L,$$

with the mesh widths (for simplicity of exposition, we assume that *mesh widths* are equal in each dimension), given by

$$(3.13) \qquad \Delta x_\ell = \Delta x(\mathcal{T}_\ell) := \frac{|I_1|}{\#\mathcal{T}_\ell^1} = \cdots = \frac{|I_d|}{\#\mathcal{T}_\ell^d} = 2^{-\ell}\Delta x_0, \quad \ell = 0, \ldots, L,$$

where Δx_0 is the mesh width for the coarsest mesh resolution $\mathcal{T}_0$ and corresponds to the lowest level $\ell = 0$, and Δx_L is the mesh width for the finest mesh resolution $\mathcal{T}_L$ and corresponds to the finest level $\ell = L$.

2. **Sample:** For each level of resolution $\ell = 0, \ldots, L$, we draw a *level-dependent* number M_ℓ of i.i.d. random input samples

$$\mathbf{I}_{\mathcal{T}_\ell}^i = \{\mathbf{U}_{0,\mathcal{T}_\ell}^i, \mathbf{S}_{\mathcal{T}_\ell}^i, (\mathbf{A}_1)_{\mathcal{T}_\ell}^i, \ldots, (\mathbf{A}_d)_{\mathcal{T}_\ell}^i\}, \quad i = 1, \ldots, M_\ell,$$

from the input random fields

$$\mathbf{I}(\omega) = \{\mathbf{U}_0(\omega), \mathbf{S}(\omega), \mathbf{A}_1, \ldots, \mathbf{A}_d\},$$

and directly approximate them by piecewise constant functions obtained from cell averaging on mesh $\mathcal{T}_\ell$ of computational domain $\mathbf{D}$, as described in subsection 3.1.

3. **Solve:** For each resolution level $\ell = 0, \ldots, L$ and for each realization of the random input data $\mathbf{I}_{\mathcal{T}_\ell}^i$ for $i = 1, ..., M_\ell$, the resulting *deterministic balance law* (1.3) (for this particular realization) is solved numerically by the Finite Volume Method (3.1) with mesh width Δx_ℓ and the corresponding

level-dependent and *realization-dependent* time step Δt_ℓ given by the CFL condition (3.2),

$$(3.14) \qquad \lambda_{\max}^i \frac{\Delta t_\ell}{\Delta x_\ell} \leq \frac{1}{2}, \quad \ell = 0, \ldots, L.$$

We denote the resulting ensemble of FVM solutions by $\mathbf{U}_{\mathcal{T}_\ell}^{i,n}$, $i = 1, \ldots, M_\ell$.

4. **Estimate solution statistics:** Firstly, we observe that the telescoping sum holds,

$$(3.15) \qquad \mathbb{E}[\mathbf{U}_{\mathcal{T}_L}] = \mathbb{E}[\mathbf{U}_{\mathcal{T}_0}] + \sum_{\ell=1}^{L} \mathbb{E}[\mathbf{U}_{\mathcal{T}_\ell} - \mathbf{U}_{\mathcal{T}_{\ell-1}}].$$

The estimate of the expectation of the random solution field is given by statistically estimating the exact mathematical expectations $\mathbb{E}[\cdot]$ for each term in the sum (3.15) using the Monte Carlo method with a *level-dependent* number of samples M_ℓ,

$$(3.16) \qquad E^L[\mathbf{U}_{\mathcal{T}_L}(\cdot, t^n)] := E_{M_0}[\mathbf{U}_{\mathcal{T}_0}^n] + \sum_{\ell=1}^{L} E_{M_\ell}[\mathbf{U}_{\mathcal{T}_\ell}^n - \mathbf{U}_{\mathcal{T}_{\ell-1}}^n],$$

with E_{M_ℓ} being the MC-FVM estimator defined in (3.8) for the mesh level $\mathcal{T}_\ell$. Higher statistical moments of $\mathbf{U}$ can be approximated analogously, we refer to [27].

The following result from [47] addresses the convergence of MLMC-FVM.

Theorem 3.4 (MLMC-FVM convergence). *Consider linear system of conservation laws (2.1) and assume that the hypotheses (A1) - (A3) of Theorem 2.5 are satisfied with $k \geq 2$, i.e. second moments of the random initial data $\mathbf{U}_0$, source $\mathbf{S}$ and K exist. Assume further that we are given a FVM such that (3.2) holds and such that Assumption 3.1 is satisfied; in particular, assume that the deterministic FVM scheme converges at rate $s > 0$ as in (3.6). Then, for time $t = t^n$, and for any sequence $\{M_\ell\}_{\ell=0}^{\infty}$ of sample sizes at mesh level ℓ, the MLMC estimate $E^L[\mathbf{U}_{\mathcal{T}}^n](\omega)$ defined in (3.16) satisfies the following error bound,*

$$(3.17)$$
$$\left\| \mathbb{E}[\mathbf{U}(t^n)] - E^L[\mathbf{U}_{\mathcal{T}_L}^n](\omega) \right\|_{L^2(\Omega, \mathbf{L}^2(\mathbf{D}))} \leq C \Delta x_L^s \|K, \mathbf{U}_0, \mathbf{S}, t^n\|_{L^1(\Omega, \mathbf{H}^s(\mathbf{D}))}$$

$$+ 2C_s \left(\sum_{\ell=1}^{L} M_\ell^{-\frac{1}{2}} \Delta x_\ell^s \right) \|K, \mathbf{U}_0, \mathbf{S}, t^n\|_{L^2(\Omega, \mathbf{H}^s(\mathbf{D}))}$$

$$+ M_0^{-\frac{1}{2}} \|K, \mathbf{U}_0, \mathbf{S}, t^n\|_{L^2(\Omega, \mathbf{L}^2(\mathbf{D}))},$$

where $C_s > 0$ is as in (3.6) and is independent of ℓ, M_ℓ and Δx_ℓ; norms are as in (2.10).

The error estimate (3.17) provided in Theorem 3.4 is the key result required to derive strategies for choosing the number of samples M_ℓ for each level $\ell = 0, \ldots, L$ in the MLMC-FVM estimator (3.17). 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 (3.17) is achieved with minimal total computational work for MLMC-FVM. We will use the asymptotically optimized number of samples derived in [11, 38], i.e. such that the *sum over all error terms* in (3.17) are *asymptotically optimized* with respect to the

required computational work. We note that such "optimization" is understood in terms of *minimizing* only the error *bound* in (3.17) and not the error directly; such error bound, of course, is only valid *asymptotically* and can be rather conservative. For each resolution level $\ell = 0, \ldots, L$, the asymptotically optimized number of samples is given by

$$(3.18) \qquad M_\ell = \left\lceil M_L 2^{\frac{2}{3}(s+d+1)(L-\ell)} \right\rceil, \quad \ell = 0, \ldots, L,$$

where $\lceil \cdot \rceil$ denotes the rounding up to integer values for number of samples and $M_L \geq 1$ denotes the number of samples on the finest level and is treated as a parameter of the MLMC-FVM algorithm. Notice that (3.18) implies that the largest number of MC samples is required on the coarsest mesh level $\ell = 0$, whereas only a small fixed number of MC samples is needed on the finest discretization levels. Choosing sample numbers M_ℓ in (3.17) according to (3.18), leads to the following error vs. work estimate for (3.16),
(3.19)

$$\left\| \mathbb{E}[\mathbf{U}(t^n)] - E^L[\mathbf{U}^n_{\mathcal{T}_L}] \right\|_{L^2(\Omega, \mathbf{L}^2(\mathbf{D}))} \lesssim \begin{cases} \text{Work}^{-s/(d+1)} & s < (d+1)/2, \\ \text{Work}^{-1/2} \log (\text{Work})^{1/2} & s = (d+1)/2, \\ \text{Work}^{-1/2} & s > (d+1)/2. \end{cases}$$

The error estimates in (3.19) show that the MLMC-FVM is superior to the MC-FVM as the asymptotic computational cost for MLMC-FVM scales as $\text{Work}^{-s/(d+1)}$; compare to $\text{Work}^{-s/(d+1+2s)}$ for the MC-FVM scheme as in (3.12). Furthermore, if $s < (d+1)/2$ then the error vs. work estimate (3.19) is of the *same order* as the error vs. work of the *deterministic* finite volume scheme, which implies that the total amount of work to achieve a certain error level ε, for instance, in approximation of the *random* entropy solution's mean field, will (asymptotically) be equal to that of approximating the entropy solution of *one deterministic* balance law at the same level L of resolution. In fact, it was shown in [44] that for the same amount of computational work the *stochastic* simulation is *at most* $\sqrt{2M_L}$ times less accurate compared to its *deterministic* version, where the free parameter M_L is usually chosen to be small, i.e. $\mathcal{O}(1) - \mathcal{O}(10)$.

We also remark, that the remaining parameters such as the number of samples on the finest mesh resolution M_L could be estimated a-priori based on several (statistical) assumptions on the (unknown) solution, refer to [44]. Several a-posteriori techniques were also proposed in [33, 38] to empirically estimate the required number of samples M_ℓ on each level as well as the optimal coarsest mesh resolution Δx_0.

4. Log-normally distributed random material coefficients

For the stochastic modeling of the uncertain material coefficient c, it is customary to assume that the Karhunen-Loève (KL) expansion [41, 47, 21] of c is available:

$$(4.1) \qquad \log c(\mathbf{x}, \omega) = \log \bar{c}(\mathbf{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m} \Psi_m(\mathbf{x}) Y_m(\omega).$$

Here, $\{\lambda_m\}_{m=1}^{\infty}$ are the eigenvalues satisfying $\{\sqrt{\lambda_m}\}_{m=1}^{\infty} \in \ell^1(\mathbb{N})$, Ψ_m are the eigenfunctions satisfying $\|\Psi_m\|_{L^2(\mathbf{D})} = 1$, $\bar{c} \in L^2(\mathbf{D})$ is the deterministic part, and Y_m are *independent* random variables with zero mean and finite variance. In

most applications of practical interest, *many* terms are needed in the Karhunen-Loève expansion to approximate the random material coefficient well. The repeated sampling of a large number of Karhunen-Loève terms is very costly, especially on the coarsest mesh levels of MLMC-FVM method, where the largest number of samples is drawn.

Hence, we propose an alternative method which allows the generation of random material coefficient samples with almost optimal (up to logarithmic terms) computational complexity on all MLMC mesh levels, including the coarsest ones.

4.1. **Spectral generator using Fast Fourier Transform.**

In the following, the spectral generator to compute log-normally distributed random field realizations based on [34, 7, 35, 42] is described. We will assume that the distribution of the random material coefficient is completely determined by the positive definite covariance operator of the random field $c(\mathbf{x}, \omega)$. Furthermore, *we assume that the covariance is stationary*, i.e. that the covariance of the (random) values at two given points $\mathbf{x}, \mathbf{y} \in \mathbf{D}$ in a *periodic* domain $\mathbf{D}$ depends only on the (component-wise) distance vector $\boldsymbol{\tau} = |\mathbf{y} - \mathbf{x}|$ between these two points,

$$(4.2) \qquad \mathrm{Cov}(\log c(\mathbf{x}, \omega), \log c(\mathbf{y}, \omega)) := k(\boldsymbol{\tau}), \quad \boldsymbol{\tau} \in \mathbf{D} = [0, p_1) \times \cdots \times [0, p_d),$$

where $k : \mathbb{R}^d \to \mathbb{R}_+$ is called an *anisotropic covariance kernel*.

For a given mesh $\mathcal{T}$, define a multi-dimensional array of cell mid-points $\mathbf{x}^{(i_1, \ldots, i_d)}$,

$$(4.3) \qquad \mathbf{X} \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d}, \quad \mathbf{X}_{i_1, \ldots, i_d} = \mathbf{x}^{(i_1, \ldots, i_d)}, \quad i_r = 1, \ldots, \#\mathcal{T}^r.$$

We are interested in approximating the values of one realization of the random material coefficient $c(\mathbf{x}, \omega)$ at the mesh discretization points $\mathbf{x}^{(i_1, \ldots, i_d)}$. We define an analogous multi-dimensional array for these values:

$$(4.4) \qquad \mathbf{c} \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d}, \quad \mathbf{c}_{i_1, \ldots, i_d} = c(\mathbf{x}^{(i_1, \ldots, i_d)}, \omega), \quad i_r = 1, \ldots, \#\mathcal{T}^r.$$

Define the covariance matrix, corresponding to the discrete version of the symmetric positive definite covariance operator, to be

$$(4.5)$$
$$\mathbf{C} \in \left(\mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d} \right)^2, \quad \mathbf{C}_{i_1, \ldots, i_d; j_1, \ldots, j_d} = \mathrm{Cov}(\log \mathbf{c}_{i_1, \ldots, i_d}; \log \mathbf{c}_{j_1, \ldots, j_d}).$$

By [43, section 4.1], the covariance matrix $\mathbf{C}$ is symmetric positive definite, hence the following decomposition of $\mathbf{C}$ exists,

$$(4.6) \qquad \mathbf{L} \in \left(\mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d} \right)^2, \quad \text{such that} \quad \mathbf{C} = \mathbf{L}\mathbf{L}^\top.$$

For instance, one choice for a matrix $\mathbf{L}$ could be the square root matrix of the matrix $\mathbf{C}$; as $\mathbf{C}$ is symmetric positive definite, $\mathbf{L}$ would also be symmetric and hence $\mathbf{L} = \mathbf{L}^\top$ would hold. Another possible choice for a matrix $\mathbf{L}$ could be the Cholesky factor $\mathbf{L}$ of the corresponding Cholesky decomposition $\mathbf{C} = \mathbf{L}\mathbf{L}^\top$.

Using the decomposition (4.6), a realization of the random material coefficient $\mathbf{c}$ can be obtained by the following steps:

(1) generate a vector $\mathbf{g}$ with i.i.d. Gaussian (standard normal) random variables as its components

$$(4.7) \qquad \mathbf{g} \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d}, \quad \mathbf{g}_{i_1, \ldots, i_d} \sim \mathcal{N}(0, 1),$$

(2) given vector $\mathbf{g}$, compute a realization of the *correlated* coefficient $\mathbf{c}$ by

$$(4.8) \qquad \mathbf{c} = \exp(\mathbf{L}\mathbf{g}),$$

where the exponential function of the vector is understood component-wise.

Then, the logarithm of the realizations of $\mathbf{c}$, generated using the above methodology, have the desired symmetric positive definite covariance matrix $\mathbf{C}$,

$$\text{Cov}(\log \mathbf{c}) = \mathbb{E}[\log \mathbf{c} \log \mathbf{c}^\top] = \mathbb{E}[\mathbf{Lg}(\mathbf{Lg})^\top] = \mathbf{L}\mathbb{E}[\mathbf{gg}^\top]\mathbf{L}^\top = \mathbf{LL}^\top = \mathbf{C}.$$

Notice that the covariance $\mathbf{C}$ is much larger than the required material coefficient array $\mathbf{c}$, hence, without further improvements, such an algorithm would be computationally extremely expensive. However, as the covariance matrix is determined by a stationary covariance kernel $k(\boldsymbol{\tau})$ on a *periodic* domain $\mathbf{D}$, see (4.2), the resulting covariance matrix $\mathbf{C}$ is *circulant*. Hence, the "matrix-vector" multiplication in (4.8) is actually a multi-dimensional "vector-vector" convolution [34, 7, 35, 42] which can be performed much faster using the Fast Fourier Transform (FFT). The algorithm has three main steps:

Step 1: d-dimensional Fourier transforms of the evaluated kernel

$$(4.9) \qquad \mathbf{k} \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d}, \quad \mathbf{k}_{i_1,\ldots,i_d} = k(|\mathbf{x}^{(i_1,\ldots,i_d)} - \mathbf{x}^{(1,\ldots,1)}|),$$

and a random Gaussian vector $\mathbf{g}$ with i.i.d. entries are needed:

$$(4.10) \qquad \hat{\mathbf{k}} = \mathcal{F}\mathbf{k} \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d}, \qquad \hat{\mathbf{g}} = \mathcal{F}\mathbf{g} \in \mathbb{C}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{C}^{\#\mathcal{T}^d}.$$

Since $\mathbf{k}$ is real and periodic, $\hat{\mathbf{k}}$ is also real and periodic (in each dimension). Moreover, the Fourier basis diagonalizes the circulant positive semi-definite matrix $\mathbf{C}$, hence $\hat{\mathbf{k}}$ is the vector of the eigenvalues of $\mathbf{C}$, i.e. all entries in $\hat{\mathbf{k}}$ are non-negative.

Step 2: Computing the square root matrix $\mathbf{L}$ as in (4.6) (with $\mathbf{L} = \mathbf{L}^\top$) corresponds to taking the element-wise square root $\hat{\mathbf{l}}$ of $\hat{\mathbf{k}}$:

$$(4.11) \qquad \hat{\mathbf{l}} \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d}, \quad \hat{\mathbf{l}}_{i_1,\ldots,i_d} = \sqrt{\hat{\mathbf{k}}_{i_1,\ldots,i_d}}.$$

Step 3: "matrix-vector" multiplication in (4.8) is equivalent to multiplying $\hat{\mathbf{l}}$ and $\hat{\mathbf{g}}$ element-wise and performing d-dimensional inverse Fourier transform $\mathcal{F}^{-1}$:

$$(4.12) \qquad \mathbf{c} = \exp(\mathcal{F}^{-1}(\hat{\mathbf{l}}\hat{\mathbf{g}})) \in \mathbb{R}^{\#\mathcal{T}^1} \times \cdots \times \mathbb{R}^{\#\mathcal{T}^d},$$

where the vector-vector multiplication and the exponential are applied element-wise. For the implementation of this generator, d-dimensional Fourier transforms $\mathcal{F}$ and $\mathcal{F}^{-1}$ were computed using the parallel version of the Fast Fourier Transform library FFTW [57].

We remark that $\hat{\mathbf{k}}$ and $\hat{\mathbf{l}}$ in steps 1 and 2 need to be computed only once for a given mesh $\mathcal{T}$, whereas $\hat{\mathbf{g}}$ in step 1 and the resulting sample $\mathbf{c}$ in step 3 need to be computed for each realization of random coefficient $\mathbf{c}$. We would also like to note that *none* of the mentioned Fourier transforms are available in complex arithmetic i.e. $\mathbb{C} \to \mathbb{C}$. In particular, since $\mathbf{k}$ is real and even, the transform $\mathbf{k} \mapsto \hat{\mathbf{k}}$ is $\mathbb{R} \to \mathbb{R}$ with $\hat{\mathbf{k}}$ even; the transform $\mathbf{g} \mapsto \hat{\mathbf{g}}$ is $\mathbb{R} \to \mathbb{C}$ with $\hat{\mathbf{g}}$ even; the final transform $\hat{\mathbf{l}}\hat{\mathbf{g}} \mapsto \mathcal{F}^{-1}(\hat{\mathbf{l}}\hat{\mathbf{g}})$ is $\mathbb{C} \to \mathbb{R}$, since both $\hat{\mathbf{l}}$ and $\hat{\mathbf{g}}$ are even, resulting in even $\hat{\mathbf{l}}\hat{\mathbf{g}}$. Performing such transforms with a standard $\mathbb{C} \to \mathbb{C}$ FFT is inefficient (though possible). In the numerical experiments reported ahead, the appropriate (hardware) *optimized* $\mathbb{R} \to \mathbb{R}$, $\mathbb{R} \to \mathbb{C}$ and $\mathbb{C} \to \mathbb{R}$ FFT transforms from the FFTW library [57] were used instead.

4.2. **Covariance upscaling.** The MLMC-FVM algorithm requires MC estimates of the *differences* $\mathbf{U}_\ell - \mathbf{U}_{\ell-1}$, i.e. solutions obtained on two consecutive mesh levels,

$$(4.13) \qquad E_{M_\ell}[\mathbf{U}_\ell - \mathbf{U}_{\ell-1}] .$$

In (4.13), the *same* realization of the random material coefficient $c(\cdot, \omega)$ is required on *different* mesh resolutions, ℓ and $\ell - 1$. The required coupling of the MLMC methods with the generation of log-normal fields using spectral FFT method is not straight-forward.

One idea was recently introduced in [34], where the coupling is achieved by drawing a realization $\mathbf{c}_\ell \in \mathbb{R}^{\#\mathcal{T}_\ell}$ of $c(,\omega)$ on the finer mesh $\mathcal{T}_\ell$, and then computing the averages of $\mathbf{c}^\ell$ on the coarser mesh to obtain the *upscaled* realization $\bar{\mathbf{c}}^{\ell-1} \in \mathbb{R}^{\#\mathcal{T}_{\ell-1}}$. One must, however, ensure that the distribution of such "averaged" discrete random fields $\bar{\mathbf{c}}^{\ell-1}(\omega)$ coincides with the distribution of the "non-averaged" discrete random fields $\mathbf{c}^{\ell-1}(\omega)$. In order to achieve this and at the same time maintain the efficiency of the method, the covariance $\mathbf{C}^\ell \in (\mathbb{R}^{\#\mathcal{T}_\ell})^2$ from level ℓ also needs to be *upscaled* to the coarser mesh resolution $\ell - 1$. In the context of stationary covariance (as described in subsection 4.1), with given kernel, only the kernel needs to be *upscaled* from $\mathbf{k}^\ell \in \mathbb{R}^{\#\mathcal{T}_\ell}$ to $\bar{\mathbf{k}}^{\ell-1} \in \mathbb{R}^{\#\mathcal{T}_{\ell-1}}$ [34]. However, in order to obtain the *upscaled* covariance kernel $\bar{\mathbf{k}}^\ell$ (and hence also $\bar{\mathbf{C}}^\ell$) for any arbitrary level $0 \le \ell < L$, a recursive computation needs to be performed starting with the covariance kernel $\mathbf{k}^L$ on the *finest* mesh resolution [34]. Such overhead in computational work would cause the MLMC algorithm to be very inefficient and would invalidate the error vs. computational work bound (3.19). Moreover, since the domain decomposition method is *not* used for samples on the coarsest mesh resolutions, the amount of memory available might be significantly smaller than is required for the computation of $\mathbf{k}^L$.

Here we present a different coupling strategy, which directly uses the spectrum $\mathbf{k}$ of the kernel k. Let $\bar{\mathbf{g}}^{\ell-1}(\omega)$ be defined as the *scaled* multi-dimensional *average* of the vector $\mathbf{g}^\ell(\omega)$,

$$(4.14) \qquad \bar{\mathbf{g}}^{\ell-1}_{i_1,\ldots,i_d}(\omega) = \frac{1}{\sqrt{2^d}} \sum_{j_r \in \{0,1\}} \mathbf{g}^\ell_{2i_1+j_1,\ldots,2i_d+j_d}(\omega), \quad i_r = 1,\ldots,\#\mathcal{T}^r_{\ell-1}.$$

The *upscaled* realization $\bar{\mathbf{c}}^{\ell-1}(\omega)$ of $\mathbf{c}^\ell$ is defined according to (4.12), using the *averaged* $\bar{\mathbf{g}}^{\ell-1}$,

$$(4.15) \qquad \bar{\mathbf{c}}^{\ell-1}(\omega) = \exp(\mathcal{F}^{-1}\hat{\mathbf{l}}^{\ell-1}\mathcal{F}\bar{\mathbf{g}}^{\ell-1}).$$

Since the (scaled) average of Gaussian i.i.d. random variables is again the same Gaussian random variable, the *upscaled* coefficient $\bar{\mathbf{c}}^{\ell-1}(\omega)$ is log-Gaussian with the same covariance kernel k. Note, that in order to have a *standard normal* distribution for $\bar{\mathbf{g}}^{\ell-1}$, the averages are multiplied by an additional scaling factor of $\sqrt{2^d}$. Moreover, the Fourier transform $\mathcal{F}\bar{\mathbf{g}}^{\ell-1}$ of the *averaged* $\mathbf{g}^\ell$ will be an *approximation* of $\mathcal{F}\mathbf{g}^\ell$, since the averaging (4.14) corresponds to the low-pass filtering of the vector $\mathbf{g}^\ell$.

In comparison to the upscaling strategy presented in [34], the above method operates only on the mesh resolution that is *one* level finer, and *not* on *every* level with finer mesh resolution.

4.3. Anisotropic exponential covariance kernel for periodic boundary conditions.
Given variance $\sigma^2 > 0$ and correlation lengths $\eta_1, \ldots, \eta_d > 0$, the exponential covariance kernel is given by

$$(4.16) \qquad k(\boldsymbol{\tau}) = \sigma^2 \exp\left(-\sqrt{\frac{\boldsymbol{\tau}_1^2}{\eta_1^2} + \cdots + \frac{\boldsymbol{\tau}_d^2}{\eta_d^2}}\right), \quad \boldsymbol{\tau} \in \mathbf{D}.$$

Since the domain $\mathbf{D} = [0, p_1) \times \cdots \times [0, p_d)$ is assumed to be periodic, the resulting kernel $k(\boldsymbol{\tau})$ is *not* smooth at hyper-planes $\boldsymbol{\tau}_r = p_r/2$, $r = 1, \ldots, d$. The exponential covariance kernel $k(\boldsymbol{\tau}(x))$ with $p_1 = 2$ and $\eta_1 = \frac{1}{2}$ is shown in Figure 1. It was observed in [34], that for long correlation lengths η_r, the resulting kernel in (4.16) defines a covariance matrix $\mathbf{C}$ that is *not* positive semi-definite [34]. Hence, following [43, section 4.2.3], we define a *periodic* exponential covariance kernel $k_P(\boldsymbol{\tau})$ by superimposing kernel (4.16) and a smooth periodic sine function,
$$(4.17)$$

$$k_P(\boldsymbol{\tau}) = \sigma^2 \exp\left(-\sqrt{\frac{(p_1 \sin(\pi/p_1 \boldsymbol{\tau}_1))^2}{\eta_1^2} + \cdots + \frac{(p_d \sin(\pi/p_d \boldsymbol{\tau}_d))^2}{\eta_d^2}}\right), \quad \boldsymbol{\tau} \in \mathbf{D}.$$

The derivation of $k_P(\boldsymbol{\tau})$ in (4.17) (as a generalization of derivation in [43, section 4.2.3]) is as follows. Define an auxiliary kernel $\bar{k}(\boldsymbol{\tau})$, obtained by tensorizing the kernel $k(\boldsymbol{\tau})$ in (4.16), resulting in a kernel defined on the $2d$-dimensional space $\mathbf{D} \times \mathbf{D}$,

$$(4.18) \qquad \bar{k}(\bar{\boldsymbol{\tau}}) = \sigma^2 \exp\left(-\sqrt{\frac{\bar{\boldsymbol{\tau}}_1^2}{\eta_1^2} + \cdots + \frac{\bar{\boldsymbol{\tau}}_{2d}^2}{\eta_{2d}^2}}\right), \quad \bar{\boldsymbol{\tau}} \in \mathbf{D} \times \mathbf{D},$$

where the correlation lengths $\eta_1, \ldots, \eta_d$ are extended to the indices $1, \ldots, 2d$ by

$$\eta_{d+1} = \eta_1, \quad \ldots, \quad \eta_{2d} = \eta_d.$$

As the kernel $(\boldsymbol{\tau})$ in (4.16) defines a positive definite covariance matrix, so does the tensorized kernel $\bar{k}(\bar{\boldsymbol{\tau}})$, but only on a tensorized domain $\mathbf{D} \times \mathbf{D}$. Then, the kernel $k_P(\boldsymbol{\tau})$ in (4.17) is obtained by restricting $\bar{k}(\bar{\boldsymbol{\tau}})$ to the d-dimensional sphere in $\mathbf{D} \times \mathbf{D}$, parametrized as follows,

$$(4.19) \quad \bar{\mathbf{x}}_r = \frac{1}{2} p_r \cdot \begin{cases} \cos(2\pi/p_r \mathbf{x}_r) & \text{if } r = 1, \ldots, d, \\ \sin(2\pi/p_{r-d} \mathbf{x}_{r-d}) & \text{if } r = d, \ldots, 2d. \end{cases} \quad \mathbf{x} \in \mathbf{D}, \quad \bar{\mathbf{x}} \in \mathbf{D} \times \mathbf{D}.$$

Using the following trigonometrical identity for each component of $\boldsymbol{\tau} = \mathbf{x} - \mathbf{x}' \in \mathbf{D}$,

$$\frac{1}{4}\left(\cos(2x_j) - \cos(2x_j')\right)^2 + \frac{1}{4}\left(\sin(2x_j) - \sin(2x_j')\right)^2 = \sin^2(x_j - x_j'), \quad j = 1, \ldots d,$$

the parametrization (4.19) of the tensorized kernel (4.18) gives the periodic kernel (4.17). Since $\bar{k}(\bar{\boldsymbol{\tau}})$ defines a positive definite covariance matrix, so does its restriction $k_P(\boldsymbol{\tau})$. Notice that the periodic exponential covariance kernel $k_P(\boldsymbol{\tau})$ depends *explicitly* on the domain geometry (periods $p_1, \ldots, p_d$) and is *smooth* with minimum at the center of $\mathbf{D}$. An example of $k_P(\boldsymbol{\tau}(x))$ for $d = 1$ with period $p_1 = 2$ and correlation length $\eta_1 = \frac{1}{2}$ is shown in Figure 1.
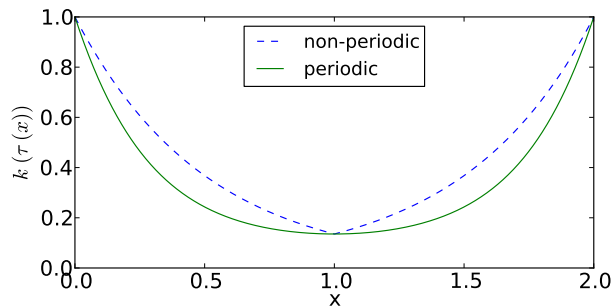
FIGURE 1. Conventional (4.16) and periodic (4.17) covariance kernels $k(\boldsymbol{\tau}(x))$ and $k_P(\boldsymbol{\tau}(x))$ for $d = 1$ with period $p_1 = 2$ and correlation length $\eta_1 = \frac{1}{2}$. The latter is smooth at $x = p_1/2$, the former is *not*.

5. Parallel implementation and adaptive run-time load balancing

The MLMC-FVM algorithm (3.16) is quite simple to implement. We remark that step 2 requires a (pseudo) random number generator. In step 3, any standard (high-order) finite volume scheme can be used. Hence, existing code for FVM can be used and there is no need to rewrite FVM code. Furthermore, the only (data) interaction between different samples is in step 4 when ensemble averages are computed. Thus, the MC-FVM is non-intrusive as well as easily parallelizable.

Our parallel implementation of MLMC-FVM scheme is called `ALSVID-UQ` [1, 46, 45]. For the detailed descriptions of how each step of the MLMC-FVM algorithm(3.16) is implemented in ALSVID-UQ, including robust pseudo random number generators (and their consistent seeding using bijective pairing functions), stable and efficient statistical estimators, parallelization paradigms and load balancing, we refer to the technical papers [46, 45]. For the wave equation with complex log-normally distributed coefficients, efficient spectral FFT generation techniques described in section 4, together with hierarchical bias-free upsacaling, were implemented using the `FFTW` library [57] for the parallelization of all required FFT transforms.

We would also like to emphasize a newly developed adaptive load balancing technique from [45], which was specifically designed to distribute very *heterogeneous* (in terms of computational work required for FVM simulations) realization-dependent MC samples. Such samples of very uneven complexity were observed in the numerical experiments (e.g. Figures 6, 10, and 15) of the forthcoming section 6. The adaptive load balancing [45] is a generalization of the greedy algorithm for workers with *heterogeneous speeds* of execution, arising due to heterogeneous domain decomposition configurations used in the parallelization of the FVM solvers. The main idea of the algorithm based on *sorting* all MC samples from all levels according to the estimated required computational run-time, and then *recursively* assigning the *largest* available sample to some (parallel) worker such that the maximal run-time of all workers (*including* the assigned sample) is *minimized*. The adaptive load balancing was shown to be a 2-approximation, i.e. the maximum run-time among all workers is at most 2 times larger than the *optimal* run-time [45]. Strong and weak

parallel scaling were achieved up to 40 000 cores on on Cray XE6 (see [58]) with 1496 AMD Interlagos 2 x 16-core 64-bit CPUs (2.1 GHz), 32 GB DDR3 memory per node, 10.4 GB/s Gemini 3D torus interconnect with a theoretical peak performance of 402 TFlops. We present the scaling results from [45] in Figure 2, where labels "MLMC" and "MLMC2" indicate that first and second order accurate FVM solvers were used.
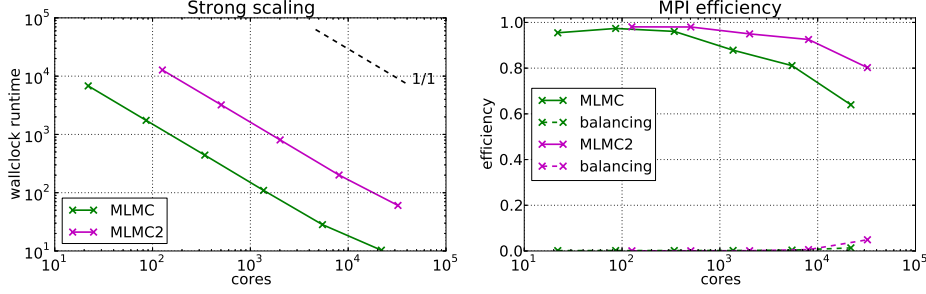

FIGURE 2. Strong scaling of the adaptive load balancing for the parallel MLMC-FVM algorithm up to 40 000 cores. The efficiency is nearly optimal, hence we expect that our parallelization algorithm will scale linearly for a much larger number of cores, especially for more computationally challenging problems.

We expect that our parallelization algorithm will scale linearly for a much larger number of cores. For the extensive description of the scaling tests and analysis of the obtained results we refer to [45].

6. ACOUSTIC ANISOTROPIC WAVE PROPAGATION IN RANDOM HETEROGENEOUS LAYERED MEDIUM

Before proceeding to the numerical experiments, we describe the notation, terminology, simulation parameter abbreviations and the methods for error computation.

6.1. **Notation, terminology and simulation parameters.** Recalling that the discretization of the random conservation law involves discretizing in space-time with a standard Finite Volume Method and the discretizing the probability space with a statistical sampling method, we tabulate various combinations of methods that are to be tested:

MC	Monte Carlo with 1st order $(s = \frac{1}{2})$ FVM	$M = \mathcal{O}(\Delta x^{-1})$,
MC2	Monte Carlo with 2nd order $(s = 1)$ FVM	$M = \mathcal{O}(\Delta x^{-2})$,
MLMC	multi-level MC with 1st order $(s = \frac{1}{2})$ FVM	$M_\ell = M_L 2^{(L-\ell)}$,
MLMC2	multi-level MC with 2nd order $(s = 1)$ FVM	$M_\ell = M_L 4^{(L-\ell)}$.

Furthermore, we need the following parameters, which will be specified for every simulation in the form of a table below the corresponding figure:

Parameter	Description
L	number of hierarchical mesh levels in the MLMC method
M_L	number of samples at the finest mesh level
grid size	number of cells in X, Y and Z directions
CFL	CFL number based on the fastest wave

cores	total number of CPU cores used in the simulation
runtime	serial runs: clock-time; parallel runs: wall-time; hrs:min:sec
efficiency	MPI efficiency, as defined in [45]

6.2. **Root mean square error estimation in MLMC-FVM.** As we will present numerical convergence analysis results, we need to specify the error estimator. In the MC-FVM and MLMC-FVM approximations (3.8) and (3.16) of the expectation $\mathbb{E}[\mathbf{U}]$, the estimators $E_M[\mathbf{U}_{\mathcal{T}}^n]$ and $E^L[\mathbf{U}_{\mathcal{T}_L}^n]$ are constructed using the *random* draws of the solution $\mathbf{U}$, and hence the estimators themselves are *random* fields, even though they are approximating *deterministic* statistical moments. For this reason, the discretization errors in (3.9) and (3.17) are random quantities as well. For our computational error convergence analysis we therefore compute a statistical estimator by averaging estimated discretization errors from several independent runs. The aforementioned errors of MC-FVM and MLMC-FVM estimators will be computed by approximating the corresponding $L^2(\Omega, \mathbf{L}^2(\mathbf{D}))$ norms with MC quadrature.

Let $E_{\mathrm{ref}}[\mathbf{U}(\mathbf{x}, t)]$ denote the reference solution, i.e. either the exact solution $\mathbb{E}[\mathbf{U}(\mathbf{x}, t)]$ (when available) or a very accurate approximation (such as on the finest resolution) of it, and

$$E_k[\mathbf{U}_{\mathcal{T}}^n(\mathbf{x})], \quad k = 1, \dots, K,$$

be a sequence of *independent* approximations of $\mathbb{E}[\mathbf{U}(\mathbf{x}, t^n)]$ obtained by running MC-FVM or MLMC-FVM solver K times, corresponding to K realizations of the stochastic space. Then the $L^2(\Omega, \mathbf{L}^2(\mathbf{D}))$-based relative error estimator is defined as in [27],

$$（6.1） \qquad \epsilon = \sqrt{\sum_{k=1}^{K} \epsilon_k^2 / K},$$

where:

$$（6.2） \qquad \epsilon_k = \frac{\|E_{\mathrm{ref}}[\mathbf{U}(\cdot, t^n)] - E_k[\mathbf{U}_{\mathcal{T}}^n(\cdot)]\|_{\mathbf{L}^2(\mathbf{D})}}{\|E_{\mathrm{ref}}[\mathbf{U}(\cdot, t^n)]\|_{\mathbf{L}^2(\mathbf{D})}} \times 100\%.$$

Definitions (6.1) - (6.2) are generalized to any E-valued statistical estimators on $\mathbf{U}$ by replacing the expectation estimators $E_{\mathrm{ref}}[\cdot]$ and $E_k[\cdot]$ with the required estimators. The extensive analysis for the appropriate choice of K is conducted in [27]; unless indicated otherwise, we choose $K = 30$ which was found to be sufficient in our numerical experiments in removing statistical fluctuations in the convergence plots.

Equipped with the above notation and concepts, we proceed to numerical experiments.

6.3. **Two dimensional case.** The physical domain is $\mathbf{D} = [0, 2] \times [0, 2]$. We assume a stationary periodic covariance kernel (4.17) with variance $\sigma^2 = 0.2$ and correlation lengths $\eta_1 = 2.0, \eta_2 = 0.1$ in (4.17). The initial data is chosen to be deterministic and is set to zero, i.e. $p_0(\mathbf{x}, \omega) \equiv 0, \mathbf{u}_0(\mathbf{x}, \omega) \equiv 0$. Identical periodic (in time) acoustic pressure pulses are injected into two locations of the domain through the *deterministic* source term f,

$$（6.3） \qquad f(\mathbf{x}, t, \omega) = \begin{cases} A_p \left(\exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_1^c\|}{2\sigma} \right) + \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_2^c\|}{2\sigma} \right) \right) & \text{if } \{3t\} < 0.02, \\ 0 & \text{else.} \end{cases}$$

where $A_p = 5000$, $\mathbf{x}_1^c = (0.5, 2.0)$, $\mathbf{x}_2^c = (1.5, 2.0)$, $\sigma^2 = 0.04$ and $\{\cdot\}$ denotes the fractional part. Perfectly reflecting boundary conditions (3.4) are assumed at the top and the bottom, and periodic boundary conditions (3.3) are assumed at the sides of the domain.

Results of a *deterministic* FVM simulation up to $t = 1.0$ are presented in Fig. 3, where the approximated sample (realization) of the random material coefficient $c(\mathbf{x}, \omega)$ and the acoustic pressure $p(\mathbf{x}, t, \omega)$ at different time instances are plotted. The computation is performed using the HLL two wave Rusanov solver and a first order accurate piecewise constant reconstruction on the mesh resolution of 4096×4096 cells, and took 7 minutes on 256 cores. Notice that since the correlation length is *long* in x-dimension and *short* in y-dimension, the resulting realization of random material coefficient $c(\mathbf{x}, \omega)$ exhibits *layered* structures as in many subsurface formations of interest: many layers with very heterogeneous wave propagation speeds are visible. The structure of the acoustic pressure wave propagation consists of close-to-circular wave fronts interfering in the center of the domain and getting distorted by the heterogeneity of the underlying physical domain.

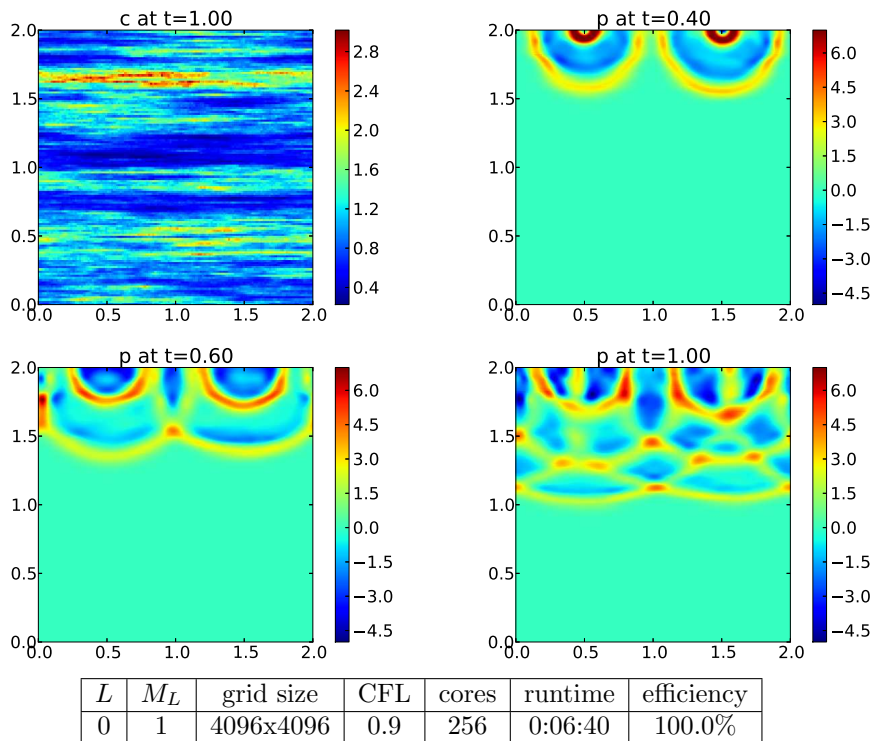| L	M_L	grid size	CFL	cores	runtime	efficiency
0	1	4096x4096	0.9	256	0:06:40	100.0%

FIGURE 3. Top left plot: one sample of the coefficient $c(x, y, \omega)$ with variance $\sigma^2 = 0.2$ and correlation lengths $\eta_1 = 2.0, \eta_2 = 0.1$. Remaining plots: time snapshots of the approximated acoustic pressure $p(x, y, \omega)$. Since the correlation length is *long* in x-dimension and *short* in y-dimension, the resulting random material coefficient $c(x, y, \omega)$ exhibits *layered* structures: notice many layers with very heterogeneous wave propagation speeds.

Results of the *stochastic* MLMC-FVM simulation up to $t = 1.0$ are presented in Figures 4 - 5, where the approximated mean and variance of the random material coefficient $c(\mathbf{x}, \omega)$ and the acoustic pressure $p(\mathbf{x}, t, \omega)$ at different time instances are provided. The computation is performed using an HLL two wave Rusanov solver and second order accurate piecewise linear WENO reconstruction. The number of levels is set to 9, i.e. $L = 8$, and the mesh resolution on the finest mesh level is 8192×8192 cells. The number of MC samples at the finest resolution is 8. We also note that the number of uncertainty sources in this simulation is very large: the material coefficient $c(\mathbf{x}, \omega)$ was sampled on the resolution of 128×128, resulting in 16 384 sources of uncertainty in the coefficients. The simulation took almost 5 hours (wall-clock) on 8176 cores (simulated on CSCS production cluster Rosa [58]).

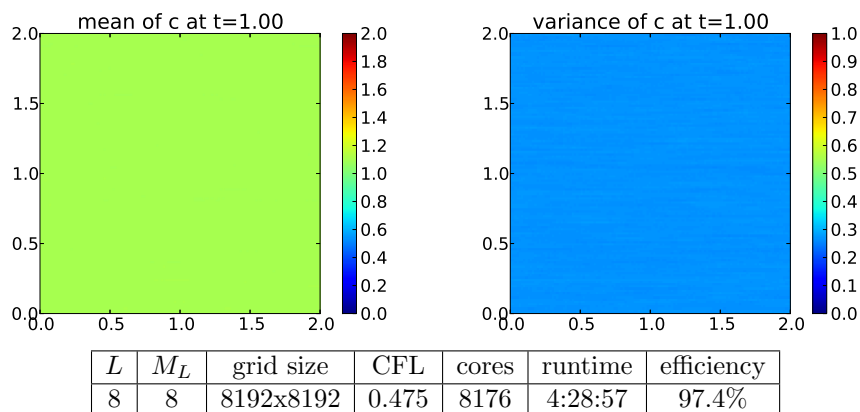| L	M_L	grid size	CFL	cores	runtime	efficiency
8	8	8192x8192	0.475	8176	4:28:57	97.4%

FIGURE 4. MLMC-FVM estimates for mean and variance of the coefficient $c(x, y, \omega)$ with variance $\sigma^2 = 0.2$ and correlation lengths $\eta_1 = 2.0, \eta_2 = 0.1$. Stationary covariance kernel resulted in homogeneous moments of the coefficient.

As expected, the mean and the variance of the material coefficient $c(\mathbf{x}, \omega)$ in Figure 4 are homogeneous within the entire domain, since a stationary covariance kernel is used. As the statistical moments of $c(\mathbf{x}, \omega)$ are known, the results in Figure 4 are also used as a "self-consistency test" of the MLMC-FVM, including the spectral FFT-based generation of the samples of the log-normally distributed material coefficient $c(\mathbf{x}, \omega)$ and the bias-free upscaling presented in section 4.

The structure of the propagation of the mean acoustic pressure waves shown in Figure 5 resembles the mean behavior of the circular interfering waves seen in the deterministic simulation of one sample, given in the previous Figure 3. The highest variance is observed at the top of the domain, i.e. at the regions where the interference of the waves is strongest.

Finally, in Figure 6, the distributions of *random* maximal wave speeds $\lambda_{\max}(\omega)$ from (3.2). across all levels are depicted, where asymptotically *optimized* numbers of samples M_ℓ as in (3.18) were used. Due to the large spread of approximately 1 to 3.5, the simulation was executed using the *adaptive* load balancing technique introduced in [45].

Next, we use the high-resolution MLMC-FVM simulations from Figure 5 as the reference solution $\mathbf{U}_{\mathrm{ref}}$ in (6.2) and investigate the convergence of the error vs.

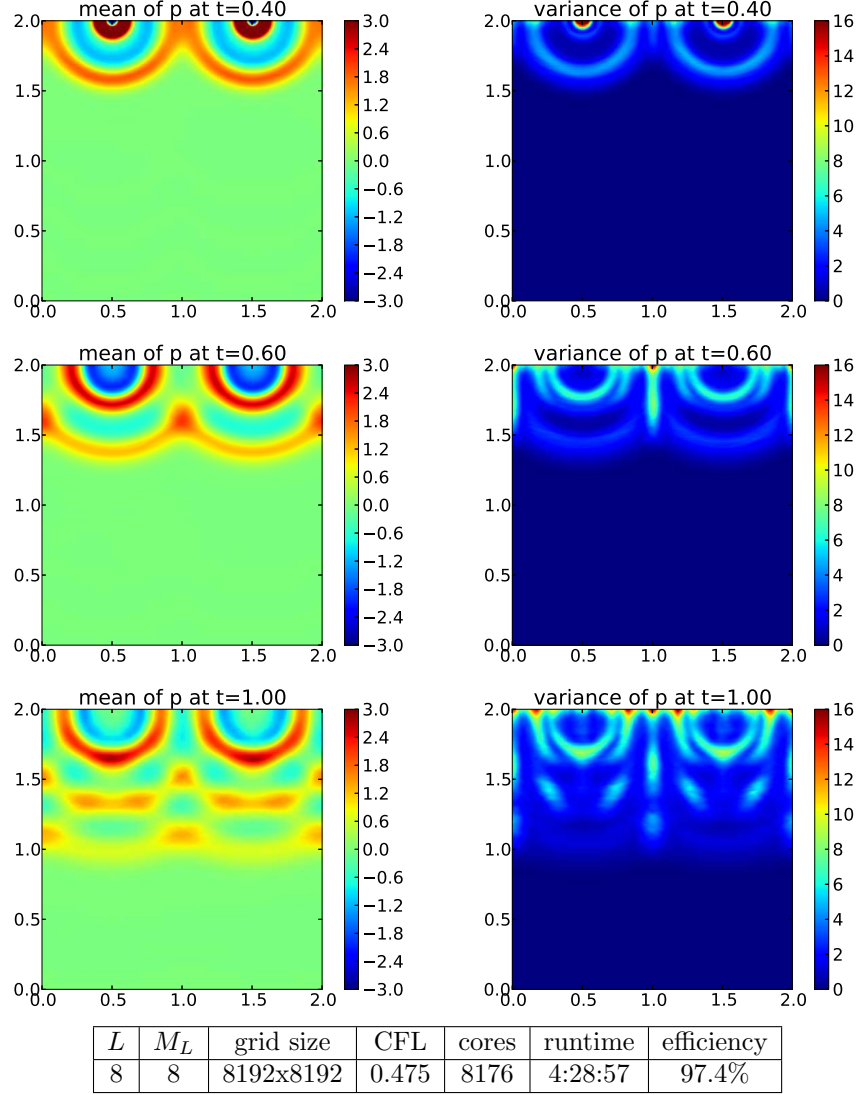| L	M_L	grid size	CFL	cores	runtime	efficiency
8	8	8192x8192	0.475	8176	4:28:57	97.4%

FIGURE 5. MLMC-FVM estimates for mean (left) and variance (right) of time snapshots of the approximated acoustic pressure $p(x, y, \omega)$. The structure of the mean acoustic pressure is *symmetric* and consists of *smooth* circular wave fronts, i.e. it resembles the mean behavior of the (distorted) circular interfering waves seen in the deterministic simulation in Figure 3. Largest variances are observed at the top of the domain.

computational work in Figure 7 and Figure 8. The errors in the mean and variance fields converge at the expected rates. At comparable numerical resolution and accuracy, we observe the MLMC(2) to be approximately one order of magnitude faster than the MC(2) method.
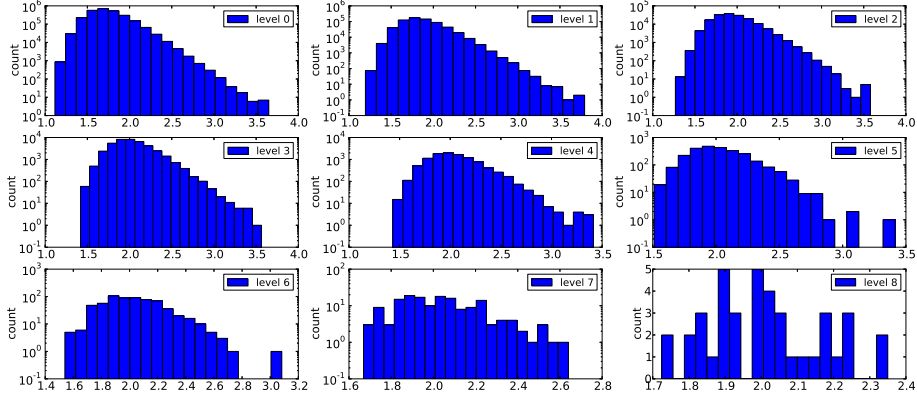
FIGURE 6. Distributions of *random* maximal wave speeds of samples for all resolution levels in the MLMC-FVM simulation reported in Figures 4 - 5. The spread is approximately from 1 to 3.5, leading to very heterogeneous computational loads for samples at the same mesh resolution level.
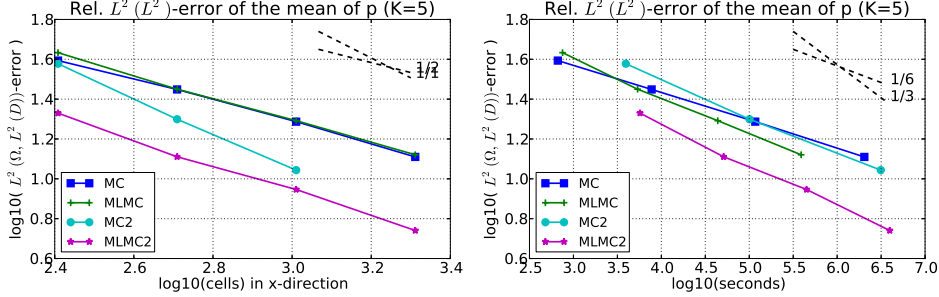

FIGURE 7. Error convergence of the mean of the acoustic pressure $p(x, y, \omega)$. The asymptotically optimized number of samples (3.18) provided slightly *higher* accuracy for the same computational work. Both MLMC2 methods are approximately two orders of magnitude faster than the MC2 method.

6.4. **Two dimensional case with discontinuous layers.** Next, in order to demonstrate the utility of the MLMC-FVM methods for very complex random material coefficients, we consider a setup analogous to subsection 6.3, but we also assume that the random material coefficient **c** is given by independent log-normal distributions within four layers D_1, D_2, D_3, D_4 partitioning the domain **D**,

$$D_1 = [0, 2] \times [0, 0.5], \quad D_2 = [0, 2] \times [0.5, 1],$$
$$D_3 = [0, 2] \times [1, 1.5], \quad D_4 = [0, 2] \times [1.5, 2].$$

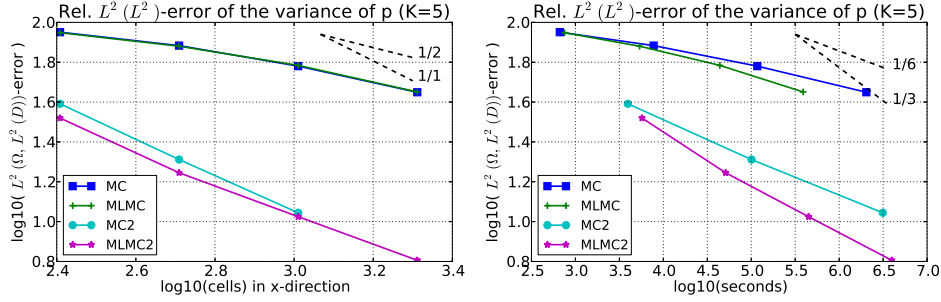FIGURE 8. Error convergence of the variance of the acoustic pressure $p(x, y, \omega)$. The asymptotically optimized number of samples (3.18) provided slightly *higher* accuracy for the same computational work. Both MLMC(2) methods are approximately one order of magnitude faster than the MC(2) method.

In each of the layers $D_1, \ldots, D_4$, we assume that the material coefficient is log-normally distributed with stationary periodic covariance kernels (4.17) and layer-dependent variances σ^2 and correlation lengths η_1, η_2, given by

$$(6.4) \qquad \sigma^2 = \begin{cases} 0.2 & \text{in } D_1 \text{ and } D_3, \\ 0.4 & \text{in } D_2 \text{ and } D_4, \end{cases} \qquad (\eta_1, \eta_2) = \begin{cases} (1.0, 0.5) & \text{in } D_1 \text{ and } D_3, \\ (2.0, 0.1) & \text{in } D_2 \text{ and } D_4. \end{cases}$$

The initial data is chosen to be deterministic and set to zero, i.e. $p_0(\mathbf{x}, \omega) \equiv 0, \mathbf{u}_0(\mathbf{x}, \omega) \equiv 0$. As in the previous experiment, identical periodic (in time) acoustic pressure pulses are injected into two locations of the domain through the *deterministic* source term f given in (6.3). Perfectly reflecting boundary conditions are assumed at the top and the bottom, and periodic boundary conditions are assumed at the sides of the domain. Results of the *deterministic* FVM simulation at $t = 1.0$ are presented in Fig. 9, where the layers $D_1, \ldots D_4$ are depicted in bottom to top order. The computation is performed using the HLL two wave Rusanov solver and a second order accurate piecewise linear WENO reconstruction on the mesh resolution of 2048×2048 cells, and took 15 minutes on 64 cores. Notice that in layers D_2 and D_4, the correlation length is *long* in x-dimension and *short* in y-dimension, resulting in even finer layered structures. On the other hand, in layers D_1 and D_3, the correlation lengths are similar in both directions, and hence no evident layered structures are present in the resulting random material coefficient. At the interfaces of layers D_1, D_2, D_3, D_4, each realization as well as the mean and variance of the random material coefficient $\mathbf{c}$ is *discontinuous*; this is expected as the distributions of $\mathbf{c}$ within each layer were assumed to be independent.

Results of the *stochastic* MLMC-FVM simulation up to $t = 1.0$ are presented in Figures 11 - 12, where the approximated mean and variance of the random material coefficient $c(\mathbf{x}, \omega)$ and the acoustic pressure $p(\mathbf{x}, t, \omega)$ at different time instances are provided. The computation is performed using the HLL two wave Rusanov solver and second order accurate piecewise linear WENO reconstruction. The number of levels is set to 5, i.e. $L = 4$, and the mesh resolution on the finest mesh level is 2048×2048 cells. The number of MC samples at the finest resolution is 16. The number of uncertainty sources is 16 384, i.e. the same as in the previous simulation.

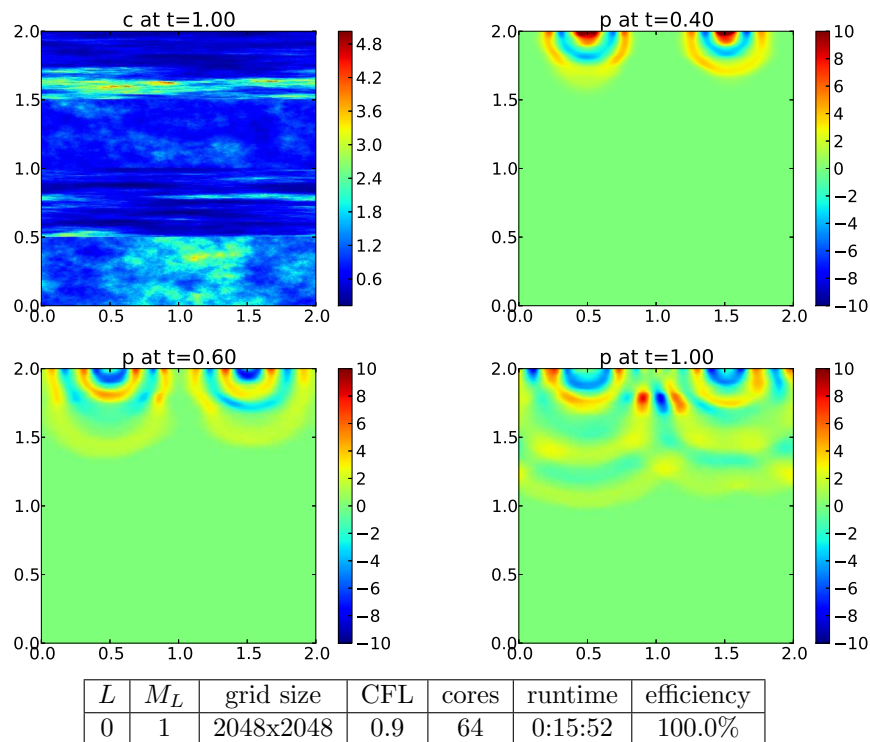| L	M_L	grid size	CFL	cores	runtime	efficiency
0	1	2048x2048	0.9	64	0:15:52	100.0%

FIGURE 9. Top left: one sample of the coefficient $c(x, y, \omega)$ with layer-dependent variances and correlation lengths as specified in (6.4). Remaining plots: time snapshots of the approximated acoustic pressure $p(x, y, \omega)$. Interchanging layers of isotropic and anisotropic material coefficient regions $c(x, y, \omega)$ are present; at the interfaces, each realization of the random material coefficient $c(x, y, \omega)$ is *discontinuous* in y-direction.

The simulation took almost 1 hour (wall-clock) on 992 cores (simulated on CSCS production cluster Rosa [58]).

The distributions of *random* maximal wave speeds across all levels are depicted in Figure 10. The spread is again large, approximately from 1.5 to 5.0, hence, *adaptive* load balancing was used, together with asymptotically *optimized* numbers of samples M_ℓ as in (3.18).

The mean and the variance of the material coefficient $c(\mathbf{x}, \omega)$ in Figure 4 consist of interchanging regions, *discontinuous* in y-direction; such heterogeneity in statistical moments is due to different covariance kernels in (6.4).

The structure of the propagation of the mean acoustic pressure waves in Figure 12 again resembles the mean behavior of the circular interfering wave seen in the deterministic simulation of one sample, depicted in Figure 9.

6.5. **Three dimensional case.** The physical domain is $\mathbf{D} = [0, 2] \times [0, 2] \times [0, 2]$. We assume a stationary periodic covariance kernel (4.17) with variance $\sigma^2 = 0.2$ and correlation lengths $\eta_1 = 2.0, \eta_2 = 0.1, \eta_3 = 2.0$ in (4.17). The initial data is chosen
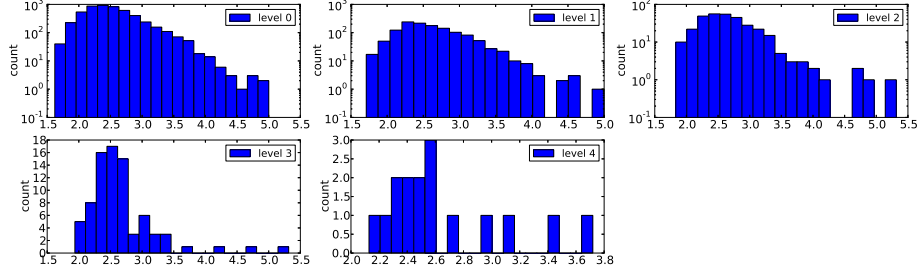
FIGURE 10. Distributions of *random* maximal wave speeds of samples for all resolution levels in the MLMC-FVM simulation reported in Figures 11 - 12. The spread is approximately from 1.5 to 5.0, leading to very heterogeneous computational loads for samples at the same mesh resolution level.

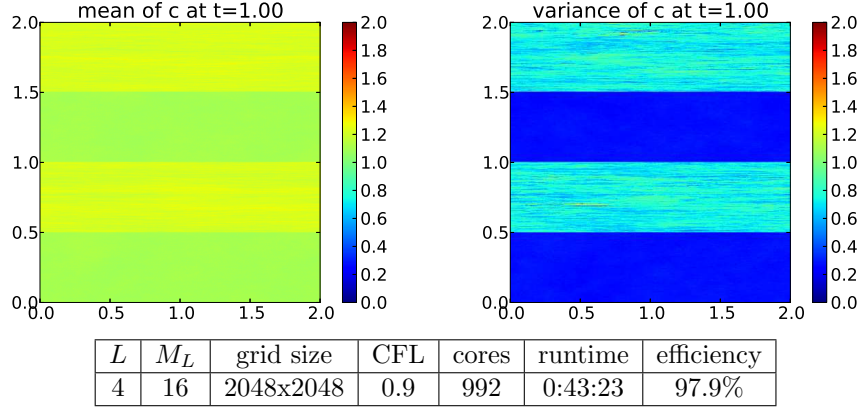| L	M_L	grid size	CFL	cores	runtime	efficiency
4	16	2048x2048	0.9	992	0:43:23	97.9%

FIGURE 11. MLMC-FVM estimates for mean and variance of the coefficient $c(x, y, \omega)$ with variances and correlation lengths as specified in (6.4). Alternating regions, *discontinuous* in y-direction, of mean and variance of the material coefficient $c(x, y, \omega)$ are due to different covariance kernels in (6.4).

to be deterministic and set to zero, i.e. $p_0(\mathbf{x}, \omega) \equiv 0, \mathbf{u}_0(\mathbf{x}, \omega) \equiv 0$. Analogously, identical periodic (in time) acoustic pressure pulses are injected into two locations of the domain through the *deterministic* source term f,

$$(6.5) \qquad f(\mathbf{x}, t, \omega) = \begin{cases} A_p \left(\exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_1^c\|}{2\sigma} \right) + \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_2^c\|}{2\sigma} \right) \right) & \text{if } \{3t\} < 0.02, \\ 0 & \text{else.} \end{cases}$$

where $A_p = 5000$, $\mathbf{x}_1^c = (0.5, 2.0)$, $\mathbf{x}_2^c = (1.5, 2.0)$, $\sigma = 0.04$ and $\{\cdot\}$ denotes the fractional part. Perfectly reflecting boundary conditions (3.4) are assumed at the top and the bottom, and the periodic boundary conditions (3.3) are assumed at the sides of the domain.

Results of the *deterministic* FVM simulation up to $t = 1.0$ are presented in Fig. 13, where the approximated sample of the random material coefficient $c(\mathbf{x}, \omega)$ and the acoustic pressure $p(\mathbf{x}, t, \omega)$ at different time instances are provided. The
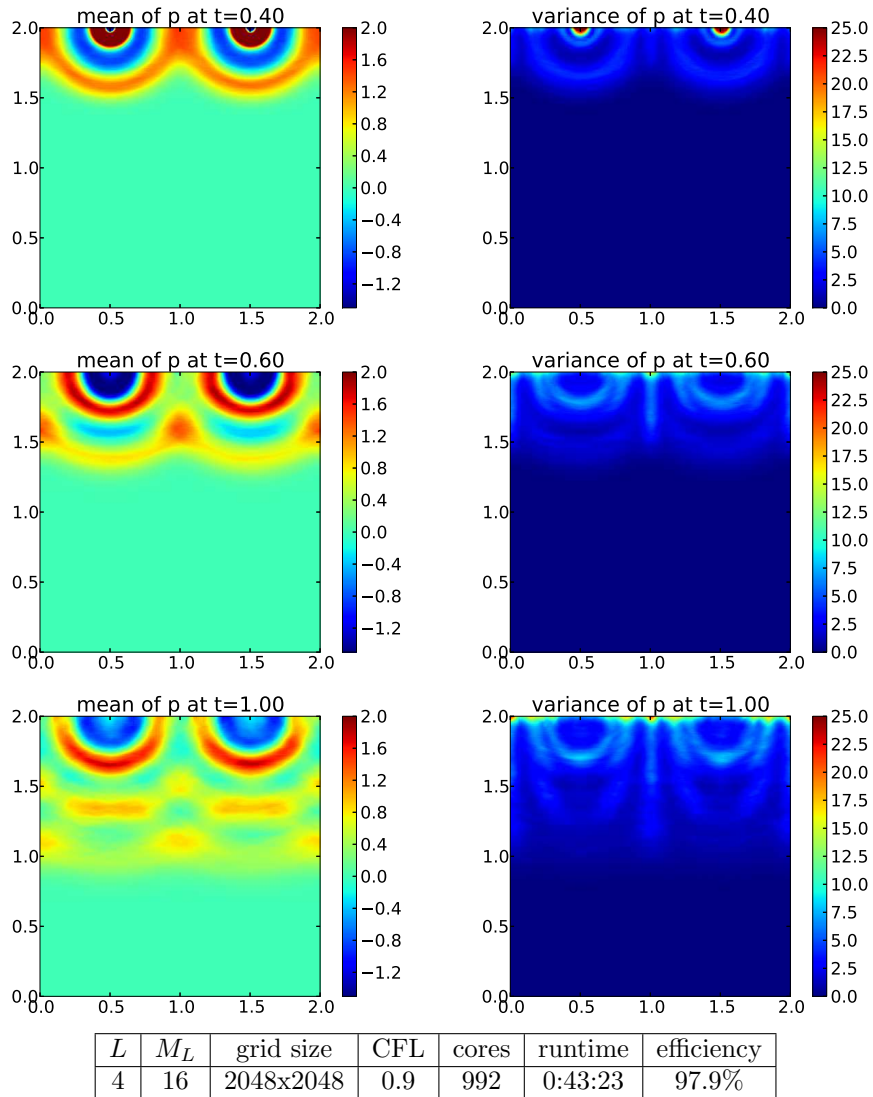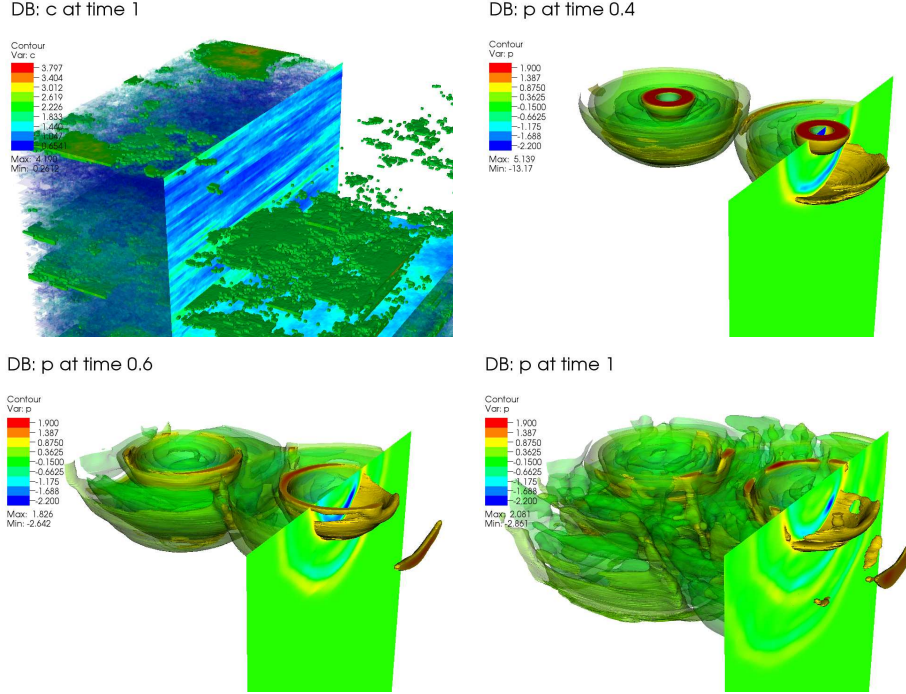
L	M_L	grid size	CFL	cores	runtime	efficiency
4	16	2048x2048	0.9	992	0:43:23	97.9%

FIGURE 12. MLMC-FVM estimates for the mean (left) and the variance (right) of time snapshots of the approximated acoustic pressure $p(x, y, \omega)$. The structure of the mean acoustic pressure is *symmetric* and consists of *smooth* circular wave fronts, i.e. it resembles the mean behavior of the (distorted) circular interfering waves seen in the deterministic simulation in Figure 9. Largest variances are observed at the top of the domain.

computation is performed using the HLL two wave Rusanov solver and a second order accurate piecewise linear WENO reconstruction on the mesh resolution of $1024 \times 1024 \times 1024$ cells, and took almost 3 hours on 4096 cores. Analogously to the two-dimensional experiment in Sect. 6.3, the *long* correlation length in xz-plane combined with *short* correlation length in y-dimension results in a random material

coefficient exhibiting *layered* structures. For a slice taken parallel to the *xz*-plane, the "entire" layer is obtained. The structure of acoustic pressure wave propagation consists of close-to-circular wave fronts interfering in the center of the domain and getting distorted by the heterogeneity of the underlying physical domain.
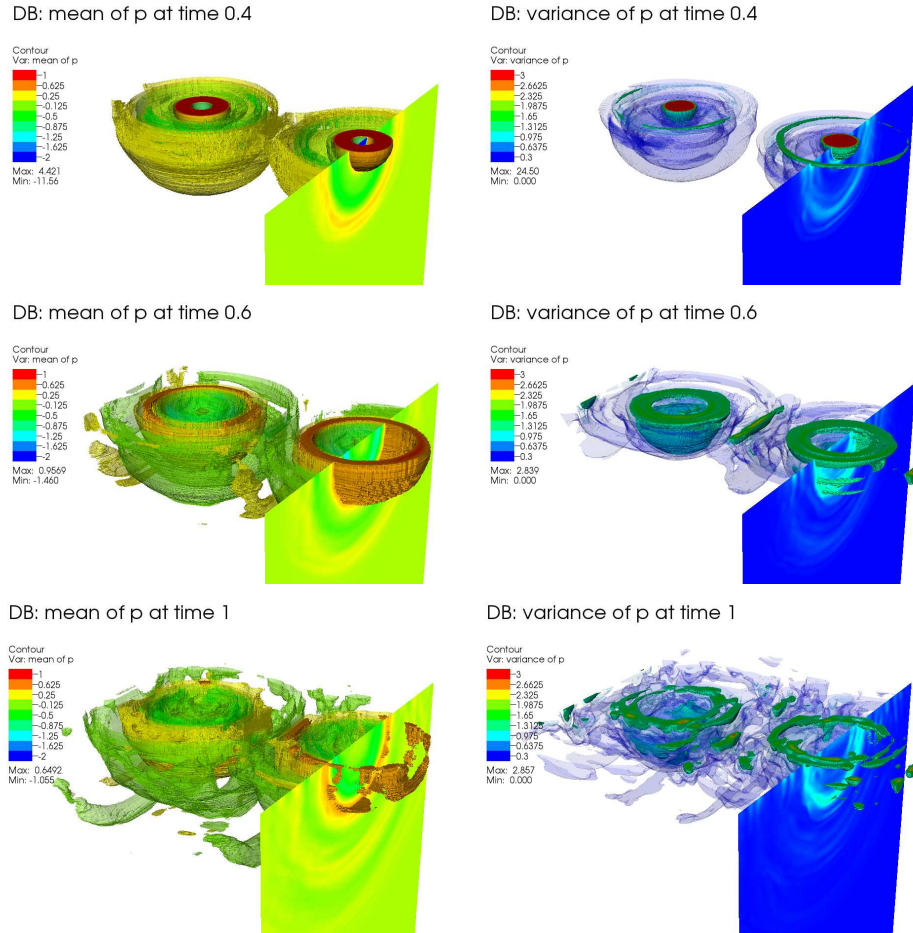

L	M_L	grid size	CFL	cores	runtime	efficiency
0	1	1024x1024x1024	0.475	4096	2:45:36	99.9%

FIGURE 13. Top left: one sample of the coefficient $c(\mathbf{x}, \omega)$ with $\sigma^2 = 0.2$ and $\eta_1 = \eta_3 = 2.0, \eta_2 = 0.1$. Remaining plots: time snapshots of the approximated acoustic pressure $p(\mathbf{x}, \omega)$. Due to anisotropic correlation lengths, coefficient $c(\mathbf{x}, \omega)$ exhibits *layered* structures, distorting the circular wave fronts of the acoustic pressure.

Results of the *stochastic* MLMC-FVM simulation up to $t = 1.0$ are presented in Figure 14, where the approximated mean and variance of the random material coefficient $c(\mathbf{x}, \omega)$ and the acoustic pressure $p(\mathbf{x}, t, \omega)$ at different time instances are provided. The computation is performed using the HLL two wave Rusanov solver and a second order accurate piecewise linear WENO reconstruction. The number of levels is set to 7, i.e. $L = 6$, and the mesh resolution on the finest mesh level is $1024 \times 1024 \times 1024$ cells. The number of MC samples at the finest resolution is 8. We also note that the number of uncertainty sources in this simulation is even larger than for the 2-D case: the material coefficient $c(\mathbf{x}, \omega)$ was sampled on the resolution of $128 \times 128 \times 128$, resulting in more than 2 *million* (2 097 152) sources

of uncertainty in coefficients. The simulation took almost 3 hours (wall-clock) on 43680 cores (simulated on CSCS production cluster Rosa [58]).

L	M_L	grid size	CFL	cores	runtime	efficiency
6	8	1024x1024x1024	0.475	43680	2:48:50	98.1%

FIGURE 14. MLMC-FVM estimates for mean (left) and variance (right) of time snapshots of the approximated acoustic pressure $p(x, y, \omega)$. The structure of the mean acoustic pressure is *symmetric* and consists of *smooth* circular wave fronts, i.e. it resembles the mean behavior of the (distorted) circular interfering waves seen in the deterministic simulation in Figure 13. Largest variances are again observed at the top of the domain.

The distributions of *random* maximal wave speeds across all levels are depicted in Figure 15. The spread is again large, approximately from 1 to 3, hence, *adaptive* load balancing was used, together with *optimal* numbers of samples M_ℓ as in (3.18).
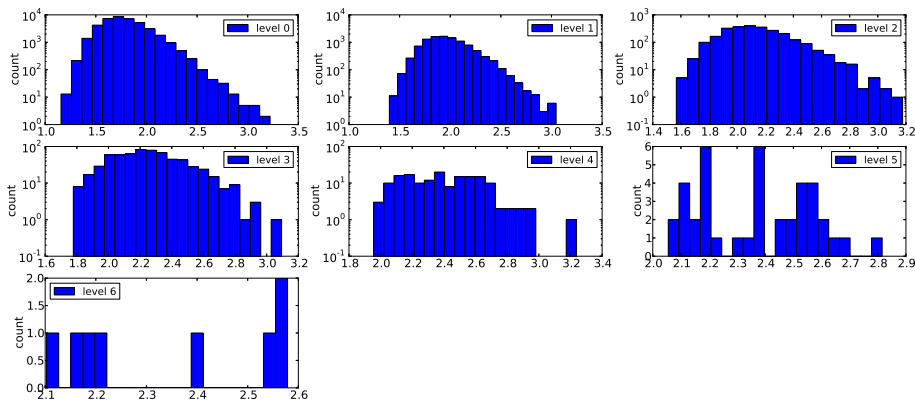
FIGURE 15. Distributions of *random* maximal wave speeds of samples for all resolution levels in the MLMC-FVM simulation reported in Figure 14. The spread is approximately from 1 to 3, leading to very heterogeneous computational loads for samples at the same mesh resolution level.

7. SUMMARY AND CONCLUSIONS

We presented a Multi-Level Monte-Carlo Finite-Volume discretization of linear, second order hyperbolic partial differential equations for the efficient computation of solution statistics for wave propagation in complex, hetergeneous random media. Specifically, we addressed the case where the acoustic properties of the heterogeneous medium are time-independent, possibly anisotropic, and are given by a lognormal gaussian random field.

Propagation of waves is simulated by explicit timestepping on a hierarchy of uniform, block-structured spatial meshes, and is subject to a CFL stability constraint.

An asymptotic convergence analysis of the total error, consisting of spatial- and temporal discretization error and Monte-Carlo sampling error was given, covering first and second order standard discretizations. Our analysis furnishes judicious choices of algorithmic steering parameters which formed the basis of and implementation on large-scale, massively parallel hardware.

In the random medium, the wave propagation speed and there also the CFL stability constraint is sample-dependent. This necessitates a novel, *probabilistic complexity and load balancing analysis* which is developed here. Due to the strongly varying, sample dependent propagation speeds, load balancing across samples become a nontrivial issue, even for solution samples on the same spatial discretization level. A novel, adaptive load-balancing scheme developed recently by one of the present authors is used to achieve near-linear scaling of the method up to several 10K processors.

A key challenge in designing efficient MLMC algorithms is the multi-scale nature of the heterogeneous, log-gaussian material properties of the random medium. A novel, bias-free multi-resolution representation of random spatially inhomogeneous coefficients was developed to ensure uniform algorithmic efficiency across all discretization levels. It is based on a novel *multi-resolution* FFT spectral sample generator with level-coupled random number generation for statistical sampling of

stochastic log-normally distributed material coefficients. In large scale numerical experiments, it is shown to maintain the efficiency of the FVM solver on coarse resolutions, which are the main building blocks for the MLMC-FVM algorithm. Numerical experiments of the acoustic wave equation with anisotropic, heterogeneous log-normally distributed material coefficients in two and three dimensions were performed, illustrating the efficiency of the MLMC-FVM method. The theoretical results and the MLMC-FVM methodology are developed for an abstract hyperbolic systems of linear balance laws, so that more complex generalizations of the acoustic wave equation can be treated with the presently proposed methods. Examples include the already mentioned elastic anisotropic wave equation with highly heterogeneous and *direction*-dependent wave propagation speeds, linearized shallow water and Euler equations.

MLMC-FVM can deal with a very large number of sources of uncertainty. For instance, the simulation of acoustic wave propagation with uncertain material coefficient involved 2 million sources of uncertainty. To the best of our knowledge, no other method (particularly deterministic methods such as quasi Monte Carlo, stochastic Galerkin, stochastic collocation, PGD, ANOVA, or stochastic FVM) is able to efficiently handle this many sources of uncertainty (i.e., high "stochastic dimensions"), in particular with solutions of low regularity and possibly non-smooth dependence on random input fields.

In the present work, we verified the strong scaling of the proposed adaptive load balancing for the MLMC-FVM method up to 40 000 cores and for more computationally challenging problems we expect it to scale linearly up to and beyond 1 million cores. However, in such large (Exa)-scale simulations on emerging massively parallel computing platforms, processor failures at run-time are inevitable and occur with increasing frequency as the number of processors increases, as reported eg. in [4]. Due to *natural fault tolerance* in MLMC-FVM due to *independent sampling*, a fault tolerant multi-level Monte Carlo (FT-MLMC) method was proposed in [36, 37, 39]. FT-MLMC does *not* rely on checkpoint/restart or on re-computation: all samples unaffected by failures are used in the computation of the final result, whereas all remaining samples affected by failures are either completely ignored or the unaffected parts in the domain decomposition parallelization of such samples are also incorporated to the final result [39]. The FT-MLMC was implemented in [37] and was shown to perform in agreement with theoretical analysis in the presence of simulated, compound Poisson distributed, random hard failures of compute cores.

We finally indicate that the presently developed multi-level methodology for the efficient computation of solution statistics is naturally suited in the context of multi-level extensions of MCMC methods for Bayesian inversion, as proposed and analyzed in [20, 21].

We also remark that, for hyperbolic, second order problems without damping or other smoothing effects, propagation of singularities occurs which precludes the use of polynomial chaos based stochastic Galerkin or collocation schemes.

Given these advantages, the Multi-level Monte Carlo Finite Volume Method, with carefully chosen number of hierarchical mesh levels and efficient implementation on massively parallel hardware architectures, appears to be a powerful general purpose technique for quantifying uncertainty in solutions of complex flow problems governed by hyperbolic systems of linear balance laws with uncertain inputs.

References

[1] ALSVID-UQ. Available from `http://www.sam.math.ethz.ch/alsvid-uq`.

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

[3] A. Barth, Ch. Schwab and N. Zollinger. *Multi-level MC Method for Elliptic PDEs with Stochastic Coefficients.* Numerische Mathematik Volume **119(1)**:123–161, 2011.

[4] F. Cappello. *Fault tolerance in petascale/exascale systems: Current knowledge, challenges and research opportunities.* Int. J. High Perform. Comput. Appl., **23(3)**:212–226, 2009.

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

[6] M. Chevreuil and A. Nouy. *Model Order Reduction based on Proper Generalized Decomposition for the Propagation of Uncertainties in Structural Dynamics.* International Journal for Numerical Methods in Engineering, **89**:241–268, 2012.

[7] J.P. Chiles and P. Delfiner, *Discrete exact simulation by Fourier method*, Geostatistics Wallagong, **96**:258–269, 1997.

[8] K.A. Cliffe, M.B. Giles, R. Scheichl and A.L. Teckentrup. *Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients*, Computing and Visualization in Science, **14(1)**:3–15, 2011.

[9] L. Evans. *Partial Differential Equations.* American Mathematical Society Providence, 1998.

[10] R. Eymard, T. Gallouët, and R. Herbin. *Finite volume methods* Handbook of numerical analysis, Vol. **VII**:713–1020, North-Holland, Amsterdam, 2000.

[11] M. Giles. *Multi-level Monte Carlo path simulation.* Oper. Res., **56**:607–617, 2008.

[12] M. Giles and Ch. Reisinger. *Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance*, Preprint July 2011, Oxford-Man Institute of Quantitative Finance and Mathematical Institute, University of Oxford.

[13] E. Godlewski and P.A. Raviart. *Hyperbolic Systems of Conservation Laws.* Mathematiques et Applications, Ellipses Publ., Paris, 1991.

[14] E. Godlewski and P.A. Raviart. *The numerical solution of multidimensional Hyperbolic Systems of Conservation Laws.* Springer Verlag Berlin Heidelberg New York, 1995.

[15] D. Gottlieb, D. Xiu. *Galerkin method for wave equations with uncertain coefficients.* Commun. Comput. Phys., **3**:505–518, 2008.

[16] S. Gottlieb, C.W. Shu and E. Tadmor. *High order time discretizations with strong stability property.* SIAM. Review, **43**:89–112, 2001.

[17] B. Gustafsson, H.O. Kreiss, and J. Oliger. *Time dependent problems and difference methods.* John Wiley & Sons, Inc., 1995.

[18] A. Harten, B. Engquist, S. Osher and S. R. Chakravarty. *Uniformly high order accurate essentially non-oscillatory schemes.* J. Comp. Phys., pp. 231–303, 1987.

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

[20] V. H. Hoang and Ch. Schwab and A. M. Stuart. *Sparse MCMC gpc Finite Element Methods for Bayesian Inverse Problems.* Inverse Problems (2014).

[21] C. Ketelsen, R. Scheichl, A.L. Teckentrup. *A Hierarchical Multilevel Markov Chain Monte Carlo Algorithm with Applications to Uncertainty Quantification in Subsurface Flow.* arXiv:1303.7343, 2013.

[22] R.A. LeVeque. *Numerical Solution of Hyperbolic Conservation Laws.* Cambridge Univ. Press 2005.

[23] G. Lin, C.H. Su and G.E. Karniadakis. *The stochastic piston problem.* Proc. Natl. Acad. Sci. USA, **101**:15840–15845, 2004.

[24] G. Lin, C.H. Su and G.E. Karniadakis. *Predicting shock dynamics in the presence of uncertainties.* J. Comp. Phys., **217**:260–276, 2006.

[25] X. Ma and N. Zabaras. *An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations.* J. Comp. Phys, **228**:3084–3113, 2009.

[26] S. Mishra, N.H. Risebro, Ch. Schwab and S. Tokareva. *Numerical solution of scalar conservation laws with random flux functions.* Submitted SIAM Journal on Uncertainty Quantification, 2014. Also available as SAM Technical Report No. 2012-35, ETH Zürich, Seminar for

Applied Mathematics, 2012, from `http://www.sam.math.ethz.ch/sam_reports/index.php?id=2012-35`.

[27] S. Mishra and Ch. Schwab. *Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data.* Math. Comp., **280**, pp. 1979–2018.

[28] S. Mishra, Ch. Schwab, and J. Šukys. *Multi-level Monte Carlo Finite Volume methods for shallow water equations with uncertain topography in multi-dimensions.* SIAM J. Sci. Comput., **34(6)**, pp. B761–B784, 2012.

[29] S. Mishra, Ch. Schwab and J. Šukys. *Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions.* J. Comput. Phys., **231(8)**, pp. 3365–3388, 2012.

[30] S. Mishra, Ch. Schwab, and J. Šukys. *Multi-level Monte Carlo finite volume methods for uncertainty quantification in nonlinear systems of balance laws.* Lecture Notes in Computational Science and Engineering Volume 92, pp. 225–294, 2013.

[31] S. Mishra and N.H. Risebro and F. Weber. *Convergence rates of finite difference schemes for the wave equation with rough coefficients.* SAM Technical Report No. 2013-42, ETH Zürich, Seminar for Applied Mathematics, 2013. Available from `http://www.sam.math.ethz.ch/sam_reports/index.php?id=2013-42`.

[32] M. Motamed, F. Nobile, and R. Tempone. *A stochastic collocation method for the second order wave equation with a discontinuous random speed*, Numer. Math., **123**:493–536, 2013.

[33] F. Müller. *Stochastic methods for uncertainty quantification in subsurface flow and transport problems.* Dissertation ETH No. 21724, 2014.

[34] F. Müller, P. Jenny, D.W. Meyer. *Multilevel Monte Carlo for two Phase Flow and Buckley-Leverett Transport in Random Heterogeneous Porous Media.* Journal of Computational Physics, available from `http://dx.doi.org/10.1016/j.jcp.2013.03.023`.

[35] E. Pardo-Iguzquiza and M. Chica-Olmo. *The Fourier integral method: an efficient spectral method for simulation of random fields.* Mathematical Geology, **25**(2):177–217, 1993.

[36] S. Pauli, P. Arbenz, and Ch. Schwab. *Intrinsic fault tolerance of Multi-level Monte Carlo methods.* SAM Technical Report No. 2012-24, ETH Zürich, Seminar for Applied Mathematics, 2012. Available from `http://www.sam.math.ethz.ch/sam_reports/index.php?id=2012-24`.

[37] S. Pauli, M. Kohler, and P. Arbenz. *A fault tolerant implementation of Multi-Level Monte Carlo methods.* Parallel Computing – Accelerating Computational Science and Engineering (CSE), Advances in Parallel Computing, IOS Press, **25:**471–480, 2014.

[38] S. Pauli and P. Arbenz. *Optimal number of multilevel Monte Carlo levels and their fault tolerant application.* Manuscript in progress, 2014.

[39] S. Pauli. *Fault tolerance in Multi-Level Monte Carlo methods.* Dissertation ETH, 2014.

[40] G. Poette, B. Després and D. Lucor. *Uncertainty quantification for systems of conservation laws.* J. Comput. Phys. **228**:2443–2467, 2009.

[41] G.D. Prato, J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, 2008.

[42] M.L. Ravalec, B. Noetinger, and L.Y. Hu. *The FFT moving average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations.* Mathematical Geology, **32**(6):701–723, 2000.

[43] C.E. Rasmussen and Ch.K.I. Williams. *Gaussian Processes for Machine Learning.* The MIT Press, 2006.

[44] J. Šukys. *Robust multi-level Monte Carlo Finite Volume methods for systems of hyperbolic conservation laws with random input data.* Dissertation ETH No. 21990, 2014.

[45] J. Šukys. *Adaptive load balancing for massively parallel multi-level Monte Carlo solvers.* PPAM 2013, Part I, LNCS 8384, pp. 47–56. Springer Berlin Heidelberg, 2014.

[46] J. Šukys, S. Mishra, and Ch. Schwab. *Static load balancing for multi-level Monte Carlo finite volume solvers.* PPAM 2011, Part I, LNCS **7203**:245–254. Springer, Heidelberg, 2012.

[47] J. Šukys, Ch. Schwab and S. Mishra. *Multi-Level Monte Carlo Finite Difference and Finite Volume methods for stochastic linear hyperbolic systems.* MCQMC 2012 Springer Proceedings in Mathematics & Statistics, Volume 65, pp. 649–666, 2013.

[48] A. L. Teckentrup. *Multilevel Monte Carlo methods for highly heterogeneous media*, Proceedings of the 2012 Winter Simulation Conference, 2012.

[49] J. Tryoen, O. Le Maitre, M. Ndjinga and A. Ern. *Intrusive projection methods with upwinding for uncertain non-linear hyperbolic systems.* J. Comput. Phys., **229**:6485–6511, 2010.

[50] J. Tryoen, O. Le Maitre, M. Ndjinga and A. Ern. *Roe solver with entropy corrector for uncertain hyperbolic systems.* J. Comp. Phys., **235**:491–506, 2010.

[51] D. Venturi, T. Sapsis, H. Cho, and G.E. Karniadakis. *A computable evolution equation for the joint response-excitation probability density function of stochastic dynamical systems.* Proceedings of the Royal Society A, **468**, 759, 2012.

[52] X. Wan and G.E. Karniadakis. *Long-term behaviour of polynomial chaos in stochastic flow simulations.* Comput. Meth. Appl. Mech. Engg. **195**:5582–5596, 2006.

[53] J.A.S. Witteveen, A. Loeven, H. Bijl *An adaptive stochastic finite element approach based on Newton-Cotes quadrature in simplex elements.* Comput. Fluids, **38**:1270–1288, 2009.

[54] J. Wloka. *Partial differential equations.* Cambridge Univ. Press, 1987.

[55] D. Xiu and J.S. Hesthaven. *High-order collocation methods for differential equations with random inputs.* SIAM J. Sci. Comput., **27**:1118–1139, 2005.

[56] X. Yang, M. Choi, G. Lin, G.E. Karniadakis. *Adaptive ANOVA decomposition of stochastic incompressible and compressible flows.* J. Comput. Phys., **231**:1587–1614, 2012.

[57] *FFTW*, available from `http://www.fftw.org/`.

[58] *Rosa (Cray XE6),* Swiss National Supercomputing Center (CSCS), Lugano, `www.cscs.ch`.

(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: `christoph.schwab@sam.math.ethz.ch`

(Jonas Šukys)
Seminar for Applied Mathematics
ETH
HG G. 62.1,
Rämistrasse 101, Zürich, Switzerland.
E-mail address: `jonas.sukys@sam.math.ethz.ch`

Recent Research Reports

Nr.	Authors/Title

2014-12   A. Paganini
Approximate Shape Gradients for Interface Problems

2014-13   E. Fonn and P. Grohs and R. Hiptmair
Polar Spectral Scheme for the Spatially Homogeneous Boltzmann Equation

2014-14   J. Dick and F.Y. Kuo and Q.T. Le Gia and Ch. Schwab
Multi-level higher order QMC Galerkin discretization for affine parametric operator equations

2014-15   Ch. Schwab
Exponential convergence of simplicial hp-FEM for H ¹-functions with isotropic singularities

2014-16   P. Grohs and S. Keiper and G. Kutyniok and M. Schaefer
α-Molecules

2014-17   A. Hiltebrand and S. Mishra
Efficient computation of all speed flows using an entropy stable shock-capturing space-time discontinuous Galerkin method

2014-18   D. Conus and A. Jentzen and R. Kurniawan
Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients

2014-19   J. Doelz and H. Harbrecht and Ch. Schwab
Covariance regularity and H-matrix approximation for rough random fields

2014-20   P. Grohs and S. Hosseini
Nonsmooth Trust Region Algorithms for Locally Lipschitz Functions on Riemannian Manifolds

2014-21   P. Grohs and A. Obermeier
Optimal Adaptive Ridgelet Schemes for Linear Transport Equations