SIAM J. NUMER. ANAL.
Vol. 52, No. 3, pp. 1243–1268

MULTILEVEL MONTE CARLO FINITE ELEMENT METHODS FOR STOCHASTIC ELLIPTIC VARIATIONAL INEQUALITIES*

RALF KORNHUBER†, CHRISTOPH SCHWAB‡, AND MAREN-WANDA WOLF†

Abstract. Multilevel Monte Carlo finite element methods (MLMC-FEMs) for the solution of stochastic elliptic variational inequalities are introduced, analyzed, and numerically investigated. Under suitable assumptions on the random diffusion coefficient, the random forcing function, and the deterministic obstacle, we prove existence and uniqueness of solutions of "pathwise" weak formulations. Suitable regularity results for deterministic, elliptic obstacle problems lead to uniform pathwise error bounds, providing optimal-order error estimates of the statistical error and upper bounds for the corresponding computational cost for the classical MC method and novel MLMC-FEMs. Utilizing suitable multigrid solvers for the occurring sample problems, in two space dimensions MLMC-FEMs then provide numerical approximations of the expectation of the random solution with the same order of efficiency as for a corresponding deterministic problem, up to logarithmic terms. Our theoretical findings are illustrated by numerical experiments.

Key words. multilevel Monte Carlo, stochastic partial differential equations, stochastic finite element methods, multilevel methods, variational inequalities

AMS subject classifications. 65N30, 65C05, 65C30, 65N55

DOI. 10.1137/130916126

1. Introduction.

Monte Carlo (MC) methods are well established in statistical simulation. For partial differential equations (PDEs) with random coefficients, numerical realization of one MC "sample" entails the numerical solution of one deterministic PDE. Many of such "paths" are required for sufficient accuracy, causing suboptimal efficiency even if optimal algebraic solvers are used (see, e.g., [5, 7, 6, 20, 31]). Multilevel versions of MC were introduced, to the best of the authors' knowledge, by Giles [21, 22] for the numerical solution of Itô stochastic ordinary differential equations, following basic ideas in an earlier work by Heinrich [28] on numerical quadrature. Such multilevel Monte Carlo (MLMC) methods have been shown to provide similar efficiency for certain stochastic PDEs as in the corresponding deterministic case [8, 34].

In the present paper, we consider elliptic obstacle problems with stochastic coefficients. Such problems arise, e.g., in the numerical simulation of subsurface flow problems or contact problems in mechanics with uncertain constitutive equations, specifically elastic moduli or friction coefficients (see, e.g., [35, 36, 38] and the references cited therein). Key characteristics of elliptic variational inequalities with stochastic coefficients are *low spatial regularity of the permeability*, *small spatial correlation lengths* (this implies slow convergence of Karhunen–Loève expansions), and the possible *nonstationarity* of realistic stochastic models, particularly from computational geosciences. All these factors hinder the efficient numerical simulation of such problems by spectral methods [19]. As for unconstrained problems, MC meth-

*Received by the editors April 8, 2013; accepted for publication (in revised form) February 18, 2014; published electronically May 29, 2014. This research was supported in part under ERC AdG grant STAHDPDE 247277.
http://www.siam.org/journals/sinum/52-3/91612.html
†FU Berlin, FB Mathematik und Informatik, Institut für Mathematik, D-14195 Berlin, Germany (kornhuber@math.fu-berlin.de, mawolf@math.fu-berlin.de).
‡Seminar für Angewandte Mathematik, ETH Zürich, CH-8092 Zürich, Switzerland (schwab@math.ethz.ch).

ods suffer from their typical lack of efficiency, even though fast multilevel solvers for discretized symmetric obstacle problems are available (see the review article [24] and the references cited therein). On this background, the present paper is devoted to the development, analysis, and implementation of multilevel Monte Carlo finite element methods (MLMC-FEMs) for symmetric, second-order, elliptic obstacle problems with random coefficients.

In this paper, the notion of randomness of diffusion coefficients is based on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with Ω denoting the set of all elementary events. We consider isotropic random diffusion coefficients $a(\cdot, \omega)$ as defined on an open, bounded Lipschitz polyhedron $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, for all $\omega \in \Omega$. These are strongly measurable mappings

$$\Omega \ni \omega \mapsto a(\cdot, \omega) \in L^\infty(D), \tag{1.1}$$

where we endow the Banach space $L^\infty(D)$ of realizations of diffusion coefficients with the sigma algebra of Borel sets to render it a measurable space. Though our algorithms will be well defined even for random coefficients whose realizations are merely in $L^\infty(D)$, we will impose stronger, spatial continuity (either $\mathbb{P}$ almost sure (a.s.) Hölder continuity or $\mathbb{P}$-a.s. continuous differentiability in $\overline{D}$) on the draws of the random coefficient in (1.1) in order to have convenient access to known regularity results for deterministic variational inequalities. For a given random source f, i.e., a strongly measurable mapping

$$\Omega \ni \omega \mapsto f(\cdot, \omega) \in L^2(D) \ , \tag{1.2}$$

and a deterministic obstacle function

$$\chi \in H^2(D), \quad \chi \leq 0 \ \text{ in } \ D \ , \tag{1.3}$$

we consider the stochastic obstacle problem which, formally in strong form, amounts to finding a random solution $u(\cdot, \omega)$ such that for $\mathbb{P}$ almost every (a.e.) $\omega \in \Omega$ and for a.e. $x \in D$ there holds that

$$\begin{aligned} -\mathrm{div}(a(\cdot,\omega)\nabla u) &\geq f(\cdot,\omega) \quad \text{in} \quad D \ , \\ u &\geq \chi \quad \text{in} \quad D \ , \\ \big(\mathrm{div}(a(\cdot,\omega)\nabla u) + f(\cdot,\omega)\big)\big(u-\chi\big) &= 0 \quad \text{in} \quad D \ , \\ u|_{\partial D} &= 0 \ . \end{aligned} \tag{1.4}$$

We concentrate on deterministic obstacle functions, because random obstacles $\chi(\cdot, \omega)$ can be traced back to the deterministic obstacle zero by introducing the new variable $w = u - \chi$. A direct treatment of stochastic obstacles is contained in [10].

The solution u of the stochastic obstacle problem (1.4) depends not only on $x \in D$ but also on the "stochastic parameter" $\omega \in \Omega$. We prove existence and uniqueness of solutions $u(\cdot, \omega)$ of "pathwise" variational (or weak) formulations of the obstacle problem (1.4), i.e., of variational formulations (with respect to $x \in D$) for $\mathbb{P}$-a.e. fixed realization or "path" $\omega \in \Omega$. We show that the collection $\{u(\omega) | \omega \in \tilde{\Omega}\}$ is measurable for some $\mathbb{P}$-measurable set $\tilde{\Omega} \in \mathcal{A}$ of full measure. In this sense, the stochastic obstacle problem (1.4) admits a unique random solution u. We show that this random solution u has finite second moments. Regularity and uniform stability of pathwise variational solutions $u(\cdot, \omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ follow from suitable regularity assumptions on the (random) data via regularity results for the deterministic, elliptic obstacle problem. From these regularity results, we obtain pathwise, optimal-order error estimates for continuous, piecewise linear finite element approximations. These error bounds hold uniformly for $\mathbb{P}$-a.e. $\omega \in \Omega$. Our arguments rely on optimal discretization error estimates for deterministic problems (see, e.g., [16, 37]), on the usual MC convergence analysis (see, e.g., [39]), and on multigrid convergence rates for the algebraic solution (see [4]). Hence, our algorithms and results readily extend to other problems with random input data for which corresponding bounds for moments, discretization errors, and convergence rates of algebraic solvers are available.

The pathwise results are the basis for the efficient computation of the expectation value of the stochastic obstacle problem (1.4) by MC-type methods. We first verify for the classical MC method the convergence order $1/2$ in terms of the number M of MC samples. Then, we show that, up to a logarithmic factor, the resulting Monte Carlo finite element method (MC-FEM) requires suboptimal computational cost of orders N^3, N^2, $N^{1+\frac{2}{3}}$ for space dimensions $d = 1, 2, 3$, respectively, in terms of the number N of degrees of freedom used in the finite element approximation. Therefore, following Barth, Schwab, and Zollinger [8], we introduce an MLMC-FEM based on a nested family of regular, simplicial triangulations of the physical domain D. In contrast to the MC-FEM, and despite the finite element spaces being nested, the MLMC-FEM does not generally preserve conformity of the corresponding deterministic sample averages for problems with inequality constraints. Assuming that suitable algebraic solvers for the pathwise sample problems are available, we show that the MLMC-FEM provides optimal-order approximations at computational cost of order N^2 for $d = 1$ and of optimal order N for problems in $d = 2, 3$ space dimensions, up to logarithmic factors.

The discretized problems which arise in sampling are discretized deterministic obstacle problems with spatially varying coefficients. They can be solved iteratively up to discretization error accuracy by recent multigrid methods [3, 4, 24, 41]. Mesh-independent polylogarithmic convergence rates, as typically observed in numerical computations, have recently been justified theoretically for so-called standard monotone multigrid (STDMMG) methods [30, 32] by Badea [4] in $d = 1, 2$ space dimensions. Hence, for $d = 2$ space dimensions, the MLMC-FEM with an algebraic STDMMG solver turns out to provide an optimal-order approximation of the expectation of the random solution u at a computational cost which is essentially the same as the cost for the numerical solution of one deterministic obstacle problem. These theoretical results are illustrated by numerical experiments in one and two space dimensions using model problems with known solutions.

The paper is organized as follows. In section 2, we collect basic properties of random fields and *elliptic variational inequalities,* which shall be used in the ensuing developments. In section 3, we state the assumptions on the random diffusion coefficient $a(\cdot, \omega)$, the random source term $f(\cdot, \omega)$, the deterministic obstacle χ, and the spatial domain D and discuss possible generalizations along with typical examples. We also provide a pathwise weak formulation of the stochastic obstacle problem (1.4) and present results on existence, uniqueness, measurability, summability, regularity, and stability of the random solution. Section 4 first addresses the convergence analysis of a stochastic finite element approximation of the pathwise variational formulation of (1.4) together with the analysis of convergence and computational cost of the MC-FEM and MLMC-FEM and then algebraic multigrid solution of the pathwise sample problems. As a corollary, we obtain almost optimal efficiency of the MLMC-FEM with an algebraic STDMMG solver for $d = 2$ space dimensions, which is one of the main results of this paper. Finally, section 5 contains several numerical experiments illustrating our theoretical findings.

2. Preliminaries.

2.1. Random fields. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, where Ω denotes a set of elementary events, $\mathcal{A} \subset 2^\Omega$ stands for the σ-algebra of all possible events, and $\mathbb{P} : \mathcal{A} \to [0, 1]$ is a probability measure. Then, for any separable Banach space X of real-valued functions on the domain $D \subset \mathbb{R}^d$ with norm $\|\cdot\|_X$, we introduce the Bochner space of strongly measurable, p-summable mappings $v : \Omega \to X$ by (see, e.g., [17, Chapter 1])

$$L^p(\Omega, \mathcal{A}, \mathbb{P}; X) := \left\{ v : \Omega \to X \mid v \text{ strongly measurable}, \quad \|v\|_{L^p(\Omega;X)} < \infty \right\},$$

where, for $0 < p \leq \infty$,

$$\|v\|_{L^p(\Omega;X)} := \begin{cases} \left(\int_\Omega \|v(\cdot, \omega)\|_X^p \, d\mathbb{P}(\omega) \right)^{1/p} & \text{if} \quad 0 < p < \infty \,, \\ \operatorname{ess\,sup}_{\omega \in \Omega} \|v(\cdot, \omega)\|_X & \text{if} \quad p = \infty \,. \end{cases}$$

In the following, we shall not explicitly indicate the dependence of Bochner spaces and their norms on the probability measure $\mathbb{P}$ if this measure is clear from the context.

Let $\mathcal{B} \in \mathcal{L}(X, Y)$ denote a continuous linear mapping from X to a separable Hilbert space Y with norm $\|\mathcal{B}\|_{X,Y}$. For a random field $v \in L^p(\Omega; X)$ this mapping defines a random field $\mathcal{B}v \in L^p(\Omega; Y)$ with the property

$$\|\mathcal{B}v\|_{L^p(\Omega;Y)} \leq \|\mathcal{B}\|_{X,Y} \|v\|_{L^p(\Omega;X)} \,.$$

Furthermore, there holds that

$$\mathcal{B} \int_\Omega v \, d\mathbb{P}(\omega) = \int_\Omega \mathcal{B}v \, d\mathbb{P}(\omega) \,.$$

We refer the reader to Chapter 1 of [17] for further results on Banach space–valued random variables.

2.2. Elliptic variational inequalities (EVIs). We briefly recall some basic existence results on deterministic EVIs, in particular the theorem of Kinderlehrer and Stampacchia [29]. Let $\mathcal{V}$ be a Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{V}}$, associated norm $\|\cdot\|_{\mathcal{V}}$, and dual space $\mathcal{V}^*$. We recall that a bilinear form $b(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is

(i) *continuous* if there exists $C_1 > 0$ such that

$$|b(u, v)| \leq C_1 \|u\|_{\mathcal{V}} \, \|v\|_{\mathcal{V}} \qquad \forall u, v \in \mathcal{V} \,, \tag{2.1}$$

(ii) *coercive* if there exists $C_2 > 0$ such that

$$b(u, u) \geq C_2 \|u\|_{\mathcal{V}}^2 \qquad \forall u \in \mathcal{V} \,. \tag{2.2}$$

THEOREM 2.1. *Let $b(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ be a continuous, coercive, not necessarily symmetric bilinear form on the Hilbert space $\mathcal{V}$, and let $\emptyset \neq \mathcal{K} \subset \mathcal{V}$ be a closed, convex subset. Then, for any $\ell \in \mathcal{V}^*$ there exists a unique solution $u \in \mathcal{K}$ of the EVI*

$$u \in \mathcal{K} : \qquad b(u, v - u) \geq \ell(v - u) \qquad \forall v \in \mathcal{K} \,. \tag{2.3}$$

In the iterative solution of discretized deterministic obstacle problems as described in section 4.5 later on, we use the following reformulation of (2.3) in terms of convex minimization that exclusively holds in the symmetric case.

Proposition 2.2. *If the bilinear form* $b(\cdot,\cdot)$ *is symmetric, then the unique solution of the EVI* (2.3) *is characterized as the unique minimizer of the associated potential*

$$J(v) := \frac{1}{2}b(v,v) - \ell(v) \ , \qquad v \in \mathcal{V} \ , \tag{2.4}$$

over the closed, convex cone $\mathcal{K}$, *i.e.,*

$$u = \arg\min \{J(v) \mid v \in \mathcal{K}\} \ . \tag{2.5}$$

For given $v \in \mathcal{V}$ the error $\|u-v\|_b$ in the energy norm $\|\cdot\|_b = b(\cdot,\cdot)^{1/2}$ and the energy error are related according to

$$\tfrac{1}{2}\|u-v\|_b^2 \le J(v) - J(u) \le \|u-v\|_b\big(\tfrac{1}{2}\|u-v\|_b + \|u-u^*\|_b\big) \ . \tag{2.6}$$

Here u^* stands for the unconstrained solution, i.e., $b(u^*,v) = \ell(v)$. Note the mismatch between the lower and upper bounds which does not occur in the unconstrained case $\mathcal{K} = \mathcal{V}$.

We shall also be interested in the following special case.

Proposition 2.3. *Assume that*

$$\mathcal{K} \subset \mathcal{V} \quad \text{is a closed, convex cone with vertex } 0 \ . \tag{2.7}$$

Then the solution $u \in \mathcal{K}$ *of the EVI* (2.3) *is characterized by*

$$u \in \mathcal{K} : \qquad b(u,v) \ge \ell(v) \quad \forall v \in \mathcal{K} \quad \text{and} \quad b(u,u) = \ell(u) \ . \tag{2.8}$$

Moreover, with the constant C_2 *as in* (2.2) *there holds the a priori estimate*

$$\|u\|_{\mathcal{V}} \le \frac{1}{C_2}\|\ell\|_{\mathcal{V}^*} \ . \tag{2.9}$$

Proof. Let $u \in \mathcal{K}$ be a solution of (2.3). As $\mathcal{K}$ is closed under linear combinations with positive coefficients, we have $w = u + v \in \mathcal{K}$ for all $v \in \mathcal{K}$. Inserting $w = u + v$ into (2.3) implies $b(u,w) \ge \ell(w)$ for all $w \in \mathcal{K}$. Inserting $w = u$ into this inequality and $v = 0 \in \mathcal{K}$ into (2.3), we get $b(u,u) = \ell(u)$. Conversely, if u solves (2.8), then we can subtract the equality from the inequality in (2.8) to show that u solves (2.3). The estimate (2.9) is a straightforward consequence of the reformulation (2.8). ∎

3. Elliptic obstacle problem with stochastic coefficients. After these preparations, we now turn to the variational formulation of the unilateral stochastic boundary value problem (1.4). To this end, we first introduce a "pathwise" abstract formulation which closely resembles the deterministic formulation (2.3) and verify its well-posedness. We then present examples of the abstract problem which, in particular, are not uniformly elliptic.

3.1. Random diffusion coefficients. We assume that the stochastic diffusion coefficient $a(x,\omega)$ is, possibly after modification on a null set, well defined and computationally accessible for *every* $\omega \in \Omega$. To ensure well-posedness later on, we impose the following assumptions on the random coefficient $a(x,\omega)$, the random source term f, and the deterministic obstacle function χ.

Assumption 3.1. The random diffusion coefficient $a(\cdot,\omega)$ and the right-hand side $f(\cdot,\omega)$ are strongly measurable mappings $\Omega \ni \omega \mapsto a(\cdot,\omega) \in L^\infty(D)$ and $\Omega \ni \omega \mapsto$

$f(\cdot,\omega) \in L^2(D)$, respectively. The random diffusion coefficient $a(\cdot,\omega)$ is elliptic in the sense that there exist real-valued random variables $\check{a}$, $\hat{a}$ such that

$$0 < \check{a}(\omega) \leq a(x,\omega) \leq \hat{a}(\omega) < \infty \qquad \text{a.e. } x \in D \tag{3.1}$$

holds for $\mathbb{P}$-a.e. $\omega \in \Omega$. We have $f(\cdot,\omega) \in L^2(D)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$, and the deterministic obstacle function $\chi \in H^2(D)$ satisfies $\chi \leq 0$ in D.

In deriving convergence rates for combined MC finite element discretizations (see Proposition 3.6 below) we will sharpen Assumption 3.1 by the following uniform ellipticity and regularity of the random coefficient $a(\cdot,\omega)$.

Assumption 3.2. (a) (uniform ellipticity) There exist constants a_-, a_+ such that the random variables $\check{a}$ and $\hat{a}$ in (3.1) satisfy $\mathbb{P}$-almost surely

$$0 < a_- \leq \check{a}(\omega) \leq \hat{a}(\omega) \leq a_+ < \infty\ , \tag{3.2}$$

and the right-hand side satisfies $f \in L^2(\Omega; L^2(D))$.

(b) (almost sure spatial regularity of a) The random diffusion coefficient $a(\cdot,\omega)$ is "pathwise" Lipschitz continuous in the sense that a is a measurable map

$$\Omega \ni \omega \mapsto a(\cdot,\omega) \in C^1(\overline{D}) \tag{3.3}$$

with the property

$$a \in L^\infty(\Omega; C^1(\overline{D}))\ . \tag{3.4}$$

(c) (regularity of D) The spatial domain $D \subset \mathbb{R}^d$ is convex.

Assumption 3.1 is satisfied, for example, for lognormal Gaussian random fields a with the choice

$$0 < \check{a}(\omega) := \operatorname*{ess\,inf}_{x\in D} a(x,\omega)\ , \qquad \hat{a}(\omega) := \operatorname*{ess\,sup}_{x\in D} a(x,\omega) < \infty\ . \tag{3.5}$$

We refer the reader to [14, Proposition 2.3], which in this case states that

$$\hat{a} \in L^p(\Omega)\ , \quad (\check{a})^{-1} \in L^p(\Omega) \qquad \forall 0 < p < \infty\ . \tag{3.6}$$

For lognormal Gaussian random fields a with sufficiently smooth covariance kernel function $R_a(\cdot,\cdot)$, the decay of the Karhunen–Loève eigenvalues to zero increases with smoothness (see, e.g., [40, Appendix]). Then, the $C^1(\overline{D})$ regularity in the sense of (3.3) in Assumption 3.2(b) is satisfied in mean square, but not $\mathbb{P}$-almost surely. Hence, we will assume occasionally that

$$a \in L^p(\Omega; C^1(\overline{D})) \qquad \forall 0 < p < \infty \tag{3.7}$$

and that the mean diffusion coefficient $\bar{a}$ satisfies

$$\bar{a} = \mathbb{E}[a] \in C^1(\overline{D})\ . \tag{3.8}$$

We emphasize that the extra assumption (Assumption 3.2) or (3.7) and (3.8) is/are imposed only to ensure $\mathbb{P}$-a.s. sufficient regularity of the random solution to yield full first-order convergence of continuous, piecewise linear finite element discretizations.

Moreover, assumption (3.3) also implies that the covariance kernel R_a of a, defined by

$$R_a := \mathbb{E}\left[(a - \mathbb{E}[a]) \otimes (a - \mathbb{E}[a])\right] \in C^1(\overline{D} \times \overline{D})\ ,$$

induces a self-adjoint integral operator $\mathcal{C}_a$, the *covariance operator*, which is compact from $L^2(D)$ to $L^2(D)$, via

$$(\mathcal{C}_a\varphi)(x) = \int_{x'\in D} R_a(x,x')\varphi(x')\,dx' \ , \quad x \in D \ .$$

The spectral theorem for compact, self-adjoint operators implies that $\mathcal{C}_a$ has a countable sequence $(\lambda_k,\varphi_k)_{k\geq 1}$ of eigenpairs with the sequence $\{\lambda_k\}_{k\geq 1}$ accumulating only at zero, and with a sequence of eigenfunctions $\varphi_k \in L^2(D)$ which we assume to be an $L^2(D)$-orthogonal, dense set in $L^2(D)$; i.e., we assume that the covariance operator $\mathcal{C}_a$ does not have finite rank or, equivalently, that the "noise" input a is genuinely infinite dimensional. Next, we present several concrete examples of random diffusion coefficients a given in terms of their Karhunen–Loève expansions.

Example 3.1 (uniform random field). Here, the random field $a \in L^2(\Omega; C^1(\overline{D}))$ is assumed to satisfy the uniform ellipticity condition (3.2); i.e., we control the random coefficient with *deterministic* lower and upper bounds a_-, a_+. The assumption $a \in L^2(\Omega; C^1(\overline{D}))$ implies the pathwise spatial regularity $a(\cdot,\omega) \in C^1(\overline{D})$ for $\mathbb{P}$-a.e. $\omega \in \Omega$, and also that $\varphi_k \in C^1(\overline{D})$. We may expand the field $a(\cdot,\omega) \in C^1(\overline{D})$ in a Karhunen–Loève series, i.e.,

$$a(x,\omega) = \bar{a}(x) + \sum_{k\geq 1} \sqrt{\lambda_k} Y_k(\omega)\varphi_k(x) \ ,$$

where, *assuming that the* φ_k *are normalized in* $L^2(D)$, the random coefficients $Y_k(\omega)$ are

$$Y_k(\omega) = \frac{1}{\sqrt{\lambda_k}} \int_{x'\in D} (a - \mathbb{E}[a])(x',\omega)\varphi_k(x')dx' \ , \quad k = 1, 2, \dots \ .$$

Example 3.2 (lognormal Gaussian random fields in D). We assume the random field a to be such that for some deterministic $\tilde{a} \in C^1(\overline{D})$ the field $g = \log(a - \tilde{a})$ is a homogeneous, Gaussian random field in D with mean $\bar{g} = \log(\bar{a}) \in C^1(\overline{D})$ and with Lipschitz continuous covariance kernel

(3.9)
$$R_g(x,x') := \mathbb{E}\left[(g(x,\cdot) - \mathbb{E}[g](x))(g(x',\cdot) - \mathbb{E}[g](x'))\right] = \rho(\|x - x'\|) \ , \quad x, \ x' \in D \ .$$

In (3.9), the function $\rho(\cdot)$ is at least Lipschitz continuous. It is well known (see, e.g., [1]) that prescribing the (deterministic) functions $\tilde{a}$, $\bar{g}$, and ρ, the stationary, Gaussian random field g is determined, up to null events.

Moreover, assuming only Lipschitz regularity of $\rho(\cdot)$ near zero, the sample paths $a(\cdot,\omega)$ belong $\mathbb{P}$-almost surely to $C^{0,s}(\overline{D})$ with $s < 1/2$ (see, e.g., [14, Proposition 2.1]). This is the case, e.g., for the so-called *exponential covariance function*. Here, ρ in (3.9) is given by $\rho_{1/2}(r) = \sigma^2 \exp(-r/\lambda)$, where $\sigma > 0$ is the variance, and $\lambda > 0$ is a correlation length parameter.

Additional smoothness of ρ near $r = 0$ implies higher spatial regularity of the realizations $a(\cdot,\omega)$. For example, for the Gaussian covariance kernel, where ρ equals $\rho_\infty(r) = \sigma^2 \exp(-r^2/\lambda^2)$, sample paths are infinite differentiable, in quadratic mean, in $\overline{D}$ (see, e.g., [1, Chapter 8]).

Both kernel functions, $\rho_{1/2}$ and ρ_∞, the exponential and Gaussian covariance kernels, are special cases of the so-called *Matérn covariances* (see, e.g., [33]), where ρ in (3.9) is given by

$$\rho_\nu(r) := \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(2\sqrt{\nu}\frac{r}{\lambda}\right)^\nu K_\nu\left(2\sqrt{\nu}\frac{r}{\lambda}\right) \ . \tag{3.10}$$

Here, K_ν denotes the modified Bessel function of the second kind. The smoothness of $\rho_\nu(\cdot)$ at $r = 0$ and, correspondingly, the spatial regularity of realizations of $a(\cdot, \omega)$ depend on the parameter ν.

3.2. Pathwise stochastic EVIs (sEVIs). For given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a separable Hilbert space V, we consider a stochastic analogue of the abstract EVI (2.3). To this end, we assume that a given random bilinear form $b(\omega; \cdot, \cdot) : V \times V \mapsto \mathbb{R}$ such that for each fixed v, $w \in V$, $\Omega \ni \omega \mapsto b(\omega; v, w)$ is a random variable which satisfies for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$|b(\omega; w, v)| \leq C_1(\omega) \|w\|_V \|v\|_V \qquad \forall v, w \in V \tag{3.11}$$

and

$$b(\omega; v, v) \geq C_2(\omega) \|v\|_V^2 \qquad \forall v \in V \tag{3.12}$$

with random variables $C_1(\omega)$, $C_2(\omega)$ with the property

$$0 < C_1(\omega) \leq C_2(\omega) < \infty \qquad \mathbb{P}\text{-a.e. } \omega \in \Omega\ . \tag{3.13}$$

In later applications, the generic conditions (3.11)–(3.13) will be ensured by Assumption 3.1, in particular by pathwise ellipticity (3.1), or by the stronger uniform ellipticity (3.2).

We further assume, given a closed, convex subset $K \subset V$ and a random linear functional, $\ell(\omega; \cdot) \in V^*$ such that, for each fixed $v \in V$, $\Omega \ni \omega \mapsto \ell(\omega; v)$ is a random variable. Then, for a given realization $\omega \in \Omega$, we consider the *"pathwise" random EVI*

$$u(\omega) \in K : \qquad b(\omega; u(\omega), v - u(\omega)) \geq \ell(\omega; v - u(\omega)) \qquad \forall v \in K\ . \tag{3.14}$$

THEOREM 3.3. *Let the assumptions* (3.11), (3.12)*, and* (3.13) *hold. Then the stochastic EVI* (3.14) *admits, for $\mathbb{P}$-a.e. $\omega \in \Omega$, a unique solution $u(\omega) \in K$. The solution correspondence*

$$\Omega \ni \omega \mapsto \Sigma(\omega) := \{u \in K | u \textit{ solves } (3.14)\} \tag{3.15}$$

is measurable with respect to the Borel σ-algebra $\mathcal{B}(V)$ of V, i.e., $u \in L^0(\Omega; V)$.

Proof. For each fixed realization $\omega \in \Omega$ the random EVI (3.14) becomes a special case of the deterministic EVI (2.3). Hence, under the assumptions (3.11), (3.12), and (3.13), existence and uniqueness of a solution $u(\omega) \in K$ of (3.14) follow from Theorem 2.1. The measurability of the solution correspondence (3.15) is shown, for example, in [26, Proposition 1.1]. ∎

We note in passing that Theorem 3.3 does not require the bilinear form $b(\omega; \cdot, \cdot)$ to be symmetric.

3.3. Pathwise variational formulation and well-posedness. Variational formulations of (1.4) will be based on the Hilbert space $V = H_0^1(D)$, which is a closed, linear subspace of $H^1(D)$. By Poincaré's inequality, the expressions

$$V \ni v \to \|v\|_V := \left(\int_D |\nabla v|^2 dx \right)^{1/2}, \quad \|v\|_a := \left(\int_D \bar{a} |\nabla v|^2 dx \right)^{1/2}$$

are equivalent norms on V. Throughout the following, we identify $L^2(D)$ with its dual and denote by V^* the dual of V with respect to the "pivot" space $L^2(D)$; i.e., we work in the triplet $V \subset L^2(D) \simeq L^2(D)^* \subset V^*$. Our MC finite element discretization will rely on a "pathwise" formulation of the form (3.14) of the stochastic elliptic obstacle problem (1.4) for a.e. ω. The set K is chosen according to

$$K := \{v \in V | v \geq \chi \text{ a.e. } x \in D\} \subset V \ , \tag{3.16}$$

with given, deterministic obstacle function χ satisfying Assumption 3.1. Note that K is a closed, convex cone with vertex $0 \in K$ (which implies (2.7)). For each realization $\omega \in \Omega$, the bilinear form $b(\omega; \cdot, \cdot)$ and right-hand side $\ell(\omega; \cdot)$ in (3.14) are given by
(3.17)

$$b(\omega; v, w) := \int_D a(x,\omega) \nabla v \cdot \nabla w \, dx \ , \quad \ell(\omega; w) := \int_D f(x,\omega) w \, dx \ , \qquad v, w \in V \ .$$

The resulting, pathwise (with respect to ω) variational (with respect to x) formulation (3.14) will be the basis of MC sampling. Well-posedness, i.e., existence, uniqueness, measurability, and stability of the solution, follow from Theorem 3.3 and Proposition 2.3.

Proposition 3.4. *Let Assumption* 3.1 *hold. Then the pathwise obstacle problem* (3.14) *admits a unique solution $u(\omega) \in K$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ which fulfills the a priori estimate*

$$\|u(\cdot, \omega)\|_V \leq \frac{1}{\check{a}(\omega)} \|f(\cdot, \omega)\|_{L^2(D)} \quad \mathbb{P}\text{-}a.e. \ \omega \in \Omega \ . \tag{3.18}$$

Imposing the stronger uniform ellipticity condition (3.2) we get the uniform estimate

$$\|u(\cdot, \omega)\|_V \leq \frac{1}{a_-} \|f(\cdot, \omega)\|_{L^2(D)} \qquad \mathbb{P}\text{-a.e. } \omega \in \Omega \ , \tag{3.19}$$

providing the moment bound $u \in L^2(\Omega; V)$ for $f \in L^2(\Omega; L^2(D))$.

Remark 3.5. From the pathwise bound (3.18), moment bounds can be obtained *without condition* (3.2). Indeed, raising (3.18) to the power $1 \leq r < \infty$ and using Hölder's inequality with conjugate indices $t, t' \geq 1$, $1/t + 1/t' = 1$ gives

$$\|u\|_{L^r(\Omega;V)} \leq \|(\check{a})^{-1}\|_{L^{rt}(\Omega)} \|f\|_{L^{rt'}(\Omega;L^2(D))} \ . \tag{3.20}$$

Hence, imposing the condition (3.6), we can select $p = rt < \infty$ sufficiently large, to find $\|u\|_{L^r(\Omega;V)} < \infty$, provided that $f \in L^{r(1+\delta)}(\Omega; L^2(D))$ for some arbitrarily small $\delta > 0$. In particular, for $r = 2$ we obtain under Assumption 3.1 that for $f \in L^{2+\delta}(\Omega; L^2(D))$ for every $\delta > 0$ there exists $C_\delta > 0$ such that

$$\|u\|_{L^2(\Omega;V)} \leq C_\delta(a) \|f\|_{L^{2+\delta}(\Omega;L^2(D))} \ . \tag{3.21}$$

3.4. Regularity of solutions. As a consequence of the regularity theory for *deterministic obstacle problems for the Laplacian* (see, e.g., [37, Chapter 5]), Assumptions 3.1 and 3.2 ensure that $u(\cdot, \omega) \in H^2(D)$ holds $\mathbb{P}$-almost surely.

Proposition 3.6. *Let Assumptions* 3.1 *and* 3.2 *hold. Then the unique pathwise solutions $u(\cdot, \omega)$ of the obstacle problems* (3.14) *belong $\mathbb{P}$-a.s. to W and the unique random solution $u \in L^2(\Omega; V)$ satisfies $L^2(\Omega; W)$. Here, the linear space W is defined by $W := \{w \in V | \Delta w \in L^2(D)\}$ and is equipped with the norm*

$$\|w\|_W := \|\Delta w\|_{L^2(D)} + \|w\|_{L^2(D)} \ .$$

Further, there holds the a priori estimate

$$\|u\|_{L^2(\Omega;W)} \leq C(a)\|f\|_{L^2(\Omega;L^2(D))} \ , \tag{3.22}$$

where $C(a)$ depends on a_- in (3.2) *and on* $\operatorname{ess\,sup}_{\omega\in\Omega} \|a(\cdot,\omega)\|_{W^{1,\infty}(D)}$.

If uniform ellipticity (3.2) *in Assumption* 3.2 *is relaxed to the (weaker) requirement* (3.6) *and the slightly stronger integrability $f \in L^{2+\delta}(\Omega; L^2(D))$ with some $\delta > 0$ of the right-hand side, then there holds that*

$$\|u\|_{L^2(\Omega;W)} \leq C_\delta(a)\|f\|_{L^{2+\delta}(\Omega;L^2(D))} \ , \tag{3.23}$$

where $C_\delta(a)$ depends on $\|(\check{a})^{-1}\|_{L^{(4+2\delta)/\delta}(\Omega)}$ and on $\operatorname{ess\,sup}_{\omega\in\Omega} \|a(\cdot,\omega)\|_{W^{1,\infty}(D)}$.

Proof. The uniform ellipticity condition (3.2) in Assumption 3.2 implies the uniform estimate (3.19). As a consequence of the $\mathbb{P}$-a.s. $W^{1,\infty}(D)$-regularity of the realizations of the stochastic coefficient $a(\omega)$ stated in Assumption 3.2, the solutions $u(\omega)$ of the pathwise variational formulation (3.14) solve the deterministic obstacle problem for the Laplacian

$$u(\omega) \in K: \qquad \int_D \nabla u(\omega) \cdot \nabla(v - u(\omega))\, dx \geq \int_D \tilde{f}(v - u(\omega))\, dx \qquad \forall v \in K \tag{3.24}$$

for $\mathbb{P}$-a.e. ω with the random source

$$\tilde{f}(\cdot,\omega) = \frac{1}{a(\cdot,\omega)} \left(f(\cdot,\omega) + \nabla a(\cdot,\omega)\cdot\nabla u(\cdot,\omega) \right) \ .$$

By Assumption 3.1, we have $f(\cdot,\omega) \in L^2(D)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$, and Assumption 3.2(b) yields uniform regularity $a(\cdot,\omega) \in C^1(\overline{D})$. Together with the a priori estimate (3.18) and uniform ellipticity $a(\cdot,\omega) \geq \check{a}(\omega) \geq a_- > 0$, as stated in Assumption 3.2(a), this implies

$$\|\tilde{f}(\cdot,\omega)\|_{L^2(D)} \leq C(a)\|f(\cdot,\omega)\|_{L^2(D)} \qquad \mathbb{P}\text{-almost everywhere in}\,\Omega \ , \tag{3.25}$$

with a constant $C(a)$ depending on a_+, a_- in (3.2) and on $\operatorname{ess\,sup}_{\omega\in\Omega} \|a(\cdot,\omega)\|_{W^{1,\infty}(D)}$ but not on $\omega \in \Omega$. As a consequence, we may estimate

$$\|\Delta u(\cdot,\omega)\|_{L^2(D)} \leq C(a)\|f(\cdot,\omega)\|_{L^2(D)} \qquad \mathbb{P}\text{-almost everywhere in}\,\Omega \ , \tag{3.26}$$

with a possibly different constant $C(a)$ independent of $\omega \in \Omega$ by utilizing convexity of D (see Assumption 3.2(c)) together with well-known $H^2(D)$ regularity results for the deterministic obstacle problem (3.24) (see, e.g., [2] or [37, Corollary 5:2.3]).

Adding the corresponding bound for $\|u(\cdot,\omega)\|_{L^2(D)}$ as implied by (3.19), squaring both sides of the resulting bound on $\|u(\cdot,\omega)\|_W$, and taking expectations implies the assertion (3.22).

If the lower bound in the uniform ellipticity (3.2) of the random coefficient is dropped, then the same reasoning as described above leads to

$$\|\tilde{f}(\cdot,\omega)\|_{L^2(D)} \leq \frac{C(a)}{\check{a}(\omega)}\|f(\cdot,\omega)\|_{L^2(D)}, \tag{3.27}$$

with $C(a)$ depending on $\operatorname{ess\,sup}_{\omega\in\Omega} \|a(\cdot,\omega)\|_{W^{1,\infty}(D)}$ but not on $\omega \in \Omega$, instead of (3.25). In the same way as described above, well-known $H^2(D)$-regularity results [37, Chapter 5] then provide the $\mathbb{P}$-a.s. bound

$$\|\Delta u(\cdot,\omega)\|_{L^2(D)} \leq \frac{C(a)}{\check{a}(\omega)}\|f(\cdot,\omega)\|_{L^2(D)}, \tag{3.28}$$

with a possibly different constant $C(a)$. Squaring (3.28) and proceeding as in Remark 3.5 with $r = 2$ and $rt' = 2 + \delta$, we obtain the desired L^2-bound (3.23), because $rt = (4 + 2\delta)/\delta$. $\square$

We recall that the first part (3.3) of Assumption 3.2(b) is satisfied, for example, for lognormal Gaussian random fields $a(\cdot, \omega)$ with sufficiently regular covariance function $R_a(x, x')$. See section 2.1 for a more detailed discussion. Note that we impose the (strong) regularity assumption, Assumption 3.2(b), only to attain full spatial regularity of solutions to the pathwise random EVIs (3.14) which in turn will provide (first-order) convergence of continuous, piecewise linear finite element discretizations later on.

Remark 3.7. Under mere Lipschitz continuity of $R_a(\cdot, \cdot)$, Assumption 3.2 does not hold, in general. In this case, only the weaker statement $\Omega \ni \omega \mapsto a(\cdot, \omega) \in C^{0,s}(\overline{D})$ with $0 < s < 1/2$ is available $\mathbb{P}$-almost surely (see [14, Proposition 1] and [15]). However, the a priori error estimates for the MC-FEM and MLMC-FEM to be stated in Theorems 4.5 and 4.9, respectively, will remain valid in this case, albeit with lower convergence rates. Note that a lack of full regularity would not affect the convergence analysis of multigrid methods for discretized pathwise EVIs (see Remark 4.13 below), but it would entail a reduced pathwise finite element convergence rate which, in turn, would lead to a rebalancing of the relation between the finite element degrees of freedom N_ℓ and the number M_ℓ of MC samples in the complexity analysis in section 4.4 ahead.

Remark 3.8. The space W can be characterized as a Sobolev space with weights vanishing at vertices and (in the case $d = 3$) at edges of the polyhedron D (see, e.g., [25]).

4. MLMC-FEM. In this section, we first introduce continuous, piecewise linear finite element discretizations of the pathwise random obstacle problems (3.14) with constraints K, bilinear form $b(\omega; \cdot, \cdot)$, and right-hand side f defined in (3.16) and (3.17), respectively. Under suitable regularity assumptions we state an optimal error estimate that holds uniformly for $\mathbb{P}$-a.e. $\omega \in \Omega$. Together with well-known convergence results on MC sampling [8] this is the main tool for the convergence and complexity analysis of the MC-FEM and MLMC-FEM for stochastic obstacle problems of the form (3.14). In the complexity analysis we assume that almost optimal algebraic solvers for the iterative solution of the discrete pathwise obstacle problems are available. Later, in section 4.5, we show that monotone multigrid (MMG) methods [4, 30, 32] have this property, at least for problems in space dimensions $d = 1, 2$. As a consequence, as in the unconstrained case, the resulting MLMC-FEM with an algebraic MMG solver (MLMC-MMG-FEM) for stochastic obstacle problems achieves almost optimal complexity.

4.1. Galerkin finite element approximation. Throughout this section we assume that D is a polyhedral domain, for simplicity. We consider a sequence of partitions $\{\mathcal{T}_l\}_{l \geq 0}$ of D into simplices resulting from uniform refinements of a coarse, regular simplicial partition $\mathcal{T}_0$ (see, e.g., [9, 11] for details). By construction, the sequence $\{\mathcal{T}_l\}_{l \geq 0}$ is shape regular (see, e.g., [12, 13]) and the mesh width h_l,

$$h_l = \max\{\mathrm{diam}(T) | T \in \mathcal{T}_l\} ,$$

of $\mathcal{T}_l$ satisfies $h_l = \frac{1}{2} h_{l-1} = 2^{-l} h_0$. For each refinement level $l = 0, 1, \ldots$ we define a corresponding sequence of nested finite element spaces

$$V_0 \subset V_1 \subset \cdots \subset V_l \subset \cdots , \tag{4.1}$$

with V_l given by

$$V_l = \{v \in V \mid v|_T \in \mathcal{P}_1 \quad \forall T \in \mathcal{T}_l\}\,. \tag{4.2}$$

Here, $\mathcal{P}_1(T) = \operatorname{span}\{x^\alpha | |\alpha| \leq 1\}$ denotes the space of linear polynomials on T so that V_l consists of all continuous functions on D, which are piecewise linear on all $T \in \mathcal{T}_l$ and satisfy homogeneous Dirichlet boundary conditions. The dimension N_l of V_l agrees with the number of elements of the set $\mathcal{N}_l$ of interior nodes of $\mathcal{T}_l$. Assuming $\chi \in H^2(D)$ (see Assumption 3.1) and utilizing $H^2(D) \subset C(\overline{D})$ for $d = 1$, 2, 3 space dimensions, we then may define

$$K_l := \{v \in V_l \mid v(p) \geq \chi(p) \ \ \forall p \in \mathcal{N}_l\}\ , \qquad \mathcal{K}_l := \overline{L^\infty(\Omega; K_l)}^{\|\cdot\|_{\mathcal{V}}}\ . \tag{4.3}$$

Note that $K_l \notin K$, with K defined in (3.16) above, and thus $\mathcal{K}_l \notin \mathcal{K}$, in general. But the sets $K_l \subset V_l$ and $\mathcal{K}_l \subset \mathcal{V}_l$ are still norm-closed, convex cones in these spaces. Under Assumption 3.1 on the obstacle χ, the cones K_l and $\mathcal{K}_l$ have common vertex 0, i.e., $0 \in K_l$ and $0 \in \mathcal{K}_l$ hold for all $l = 0, 1, 2, \ldots$ so that Proposition 2.3 is applicable.

The discretization of weak form (3.14) of the stochastic elliptic boundary value problem (1.4) reads as follows: for $\omega \in \Omega$, given $f(\omega) \in L^2(D)$, find

$$u_l(\omega) \in K_l : \quad b(\omega; u_l(\omega), v_l - u_l(\omega)) \geq \ell(v_l - u_l(\omega)) \quad \forall v_l \in K_l\,, \qquad \omega \in \Omega\,, \tag{4.4}$$

of the pathwise sEVI (3.14). In particular, for each draw of $a(\cdot, \omega)$, a discrete deterministic obstacle problem of the form (4.4) will have to be solved in both MC and MLMC sampling. As in Proposition 3.4, existence, uniqueness, and a priori estimates and moment bounds follow from Theorem 3.3 and Proposition 2.3, respectively.

PROPOSITION 4.1. *Let Assumption* 3.1 *hold. Then there exists a unique solution* $u_l(\omega) \in K_l$ *of* (4.4) *for* $\mathbb{P}$*-a.e.* $\omega \in \Omega$ *and* $l \in \mathbb{N}$ *that satisfies*

$$\sup_{l\in\mathbb{N}} \|u_l(\omega)\|_V \leq \frac{1}{\check{a}(\omega)} \|f(\omega)\|_{L^2(D)} \qquad \mathbb{P}\text{-}a.e.\ \omega \in \Omega\,. \tag{4.5}$$

Under the additional assumption, Assumption 3.2(a)*, the family of pathwise solutions* $\{u_l(\omega) | \omega \in \Omega\} \subset L^\infty(\Omega; K)$ *satisfies the a priori estimate*

$$\sup_{l\in\mathbb{N}} \|u_l(\omega)\|_V \leq \frac{1}{a_-} \|f(\omega)\|_{L^2(D)} \qquad \mathbb{P}\text{-}a.e.\ \omega \in \Omega\,. \tag{4.6}$$

Now we are ready to show a pathwise error estimate that holds uniformly $\mathbb{P}$-almost everywhere in Ω.

PROPOSITION 4.2. *Let Assumptions* 3.1 *and* 3.2 *hold. Then there exists a positive constant* $C = C(a, f, \chi)$*, independent of* $\omega \in \Omega$ *and* $l \in \mathbb{N}$*, such that*

$$\|u(\omega) - u_l(\omega)\|_V \leq C h_l \qquad \mathbb{P}\text{-}a.e.\ \omega \in \Omega\,. \tag{4.7}$$

Proof. We define the representation $A(\omega) : V \to V^*$ of $b(\omega; \cdot, \cdot)$ via the Riesz representation theorem by $\langle A(\omega)w, v\rangle = b(\omega; w, v)$ for all $v, w \in V$, where $\langle \cdot, \cdot \rangle$ denotes the duality pairing of V^* and V. By Proposition 3.6, Assumptions 3.1 and 3.2 imply $H^2(D)$-regularity, i.e.,

$$A(\omega)u(\omega) = -\nabla \cdot (a(\omega)\nabla u(\omega)) \in L^2(D)$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$. Then the pathwise bound

$$
\begin{aligned}
&\|u(\omega) - u_l(\omega)\|_V^2 \\
&\le \inf_{v_l \in K_l} \left(\frac{2}{\check{a}(\omega)} \|f(\omega) - A(\omega)u(\omega)\|_{L^2(D)} \|u(\omega) - v_l\|_{L^2(D)} + \frac{\hat{a}(\omega)^2}{\check{a}(\omega)^2} \|u(\omega) - v_l\|_V^2 \right) \\
&\qquad\qquad + \|f(\omega) - A(\omega)u(\omega)\|_{L^2(D)} \inf_{v \in K} \|u_l(\omega) - v\|_{L^2(D)} \\
&= I + II
\end{aligned}
\tag{4.8}
$$

follows directly from Theorem 1 in the classical paper of Falk [18] (see also, e.g., [16, Theorem 5.1.1], in particular [16, equation (5.1.11)]). We estimate the terms I and II separately. The error estimate

$$
I \le C(a, f) h_l^2 \qquad \mathbb{P}\text{-a.e. } \omega \in \Omega \tag{4.9}
$$

follows from the regularity assumption (3.3), the a priori bound (3.26), uniform ellipticity (3.2), and well-known interpolation error estimates (see, e.g., [16, Theorem 3.1.6]). The remaining deterministic estimate

$$
II \le C(\chi) h_l^2
$$

is stated in the proof of [16, Theorem 5.1.2]. $\square$

We proceed with an analysis of the rate of convergence of the MC method for the solution of the stochastic elliptic problem (3.14). First we derive the estimate for the solution which is not discretized in space and then generalize this result to the finite element solution.

4.2. Rate of convergence of the MC method. We estimate the expectation $\mathbb{E}[u] \in V$ by the sample average $E_M[u]$,

$$
E_M[u] := \frac{1}{M} \sum_{i=1}^{M} u^i \in V\ , \tag{4.10}
$$

over solution samples $u^i \in V$, $i = 1, \dots, M$, corresponding to M independent, identically distributed realizations of the random input data a and f.

The following result is a bound on the *statistical error* resulting from this MC estimator.

LEMMA 4.3. *For any $M \in \mathbb{N}$ and for all $u \in L^2(\Omega; V)$ we have the error estimate*

$$
\|\mathbb{E}[u] - E_M[u]\|_{L^2(\Omega;V)} \le M^{-1/2} \|u\|_{L^2(\Omega;V)}\ . \tag{4.11}
$$

Proof. With the usual interpretation of the sample average $E_M[u]$ as a V-valued random variable, the independence of the identically distributed samples u^i implies

$$
\begin{aligned}
\|\mathbb{E}[u] - E_M[u]\|_{L^2(\Omega;V)}^2 &= \mathbb{E}\left[\left\| \mathbb{E}[u] - \frac{1}{M} \sum_{i=1}^{M} u^i \right\|_V^2 \right] = \frac{1}{M^2} \sum_{i=1}^{M} \mathbb{E}\left[\|\mathbb{E}[u] - u^i\|_V^2 \right] \\
&= \frac{1}{M} \mathbb{E}\left[\|\mathbb{E}[u] - u\|_V^2 \right] = \frac{1}{M} (\mathbb{E}\|u\|_V^2 - \|\mathbb{E}[u]\|_V^2) \\
&\le \frac{1}{M} \|u\|_{L^2(\Omega;V)}^2\ . \qquad \square
\end{aligned}
$$

4.3. Single-level MC-FEM. MC-FEMs are obtained by suitable finite element approximations of the "samples" u^i occurring in (4.10). To this end, we replace u^i by Galerkin finite element approximations $u_l^i = u_l(\omega^i) \in K_l$ that can be computed from the discrete deterministic obstacle problems (4.4) with corresponding ω^i. The Monte Carlo finite element (MC-FE) approximation of $\mathbb{E}[u]$ thus reads as

$$E_M[u_l] := \frac{1}{M}\sum_{i=1}^{M} u_l^i \in L^2(\Omega; V_l)\,. \tag{4.12}$$

Remark 4.4. The MC-FE approximation is conforming in the sense that $E_M[u_l] \in \mathcal{K}_l$ holds for all $M \in \mathbb{N}$, because K_l is convex, and $E_M[u_l]$ is a convex combination of elements $u_l^i \in K_l$.

We now establish a first error estimate for the MC-FEM.

THEOREM 4.5. *Let Assumptions* 3.1 *and* 3.2 *hold. Then we have the error bound*

$$\|\mathbb{E}[u] - E_M[u_l]\|_{L^2(\Omega;V)} \leq C(a,f,\chi)\left(M^{-\frac{1}{2}} + h_l\right)\,. \tag{4.13}$$

Proof. We split the left-hand side of the above estimate as follows:

$$\begin{aligned}\|\mathbb{E}[u] - E_M[u_l]\|_{L^2(\Omega;V)} &\leq \|\mathbb{E}[u] - \mathbb{E}[u_l]\|_{L^2(\Omega;V)} + \|\mathbb{E}[u_l] - E_M[u_l]\|_{L^2(\Omega;V)} \\ &\leq \mathbb{E}\left[\|u - u_l\|_V\right] + \|\mathbb{E}[u_l] - E_M[u_l]\|_{L^2(\Omega;V)}\,.\end{aligned}$$

The first term on the right-hand side is bounded according to Proposition 4.2. Utilizing Lemma 4.3, the second term is bounded by $M^{-\frac{1}{2}}\|u_l\|_{L^2(\Omega;V)}$, and Proposition 4.2 implies that $\|u_l\|_{L^2(\Omega;V)}$ is bounded by $Ch_l + \|u\|_{L^2(\Omega;V)}$ with $C = C(a,f,\chi)$ independent of l. This completes the proof. ◻

A choice of sample size M versus grid size h_l to reach a prescribed error level is obtained by equilibrating the statistical and the discretization error in (4.13). Hence, Theorem 4.5 yields

$$M = \mathrm{O}(h_l^{-2}) = \mathrm{O}(2^{2l})\,. \tag{4.14}$$

We now provide an upper bound for the computational cost of the MC-FEM (4.12) under the assumption that discrete obstacle problems of the form (4.4) can be solved up to discretization accuracy in near optimal complexity, which will be verified in what follows.

Assumption 4.6. Approximations $\tilde{u}_l(\omega)$ of solutions $u_l(\omega)$ of deterministic obstacle problems of the form (4.4) which provide the error bound

$$\|u_l(\omega) - \tilde{u}_l(\omega)\|_V \leq Ch_l\,, \qquad l = 0, 1, \dots\,,$$

can be evaluated at computational cost which is bounded, as $l \to \infty$, by $\mathrm{O}((1+l)^\mu N_l)$ with some nonnegative integer μ and the constant implied in $\mathrm{O}(\cdot)$ being independent of l.

THEOREM 4.7. *Consider some fixed* $L \in \mathbb{N}$, *and let Assumptions* 3.1*,* 3.2*, and* 4.6 *hold. Then the inexact MC-FE approximation*

$$E_M[\tilde{u}_L] = \frac{1}{M}\sum_{i=1}^{M} \tilde{u}_L^i\,, \quad M = \mathrm{O}(2^{2L})\,,$$

of $\mathbb{E}[u]$ *with accuracy*

$$\|\mathbb{E}[u] - E_M[\tilde{u}_L]\|_{L^2(\Omega;V)} = \mathrm{O}(h_L)$$

can be evaluated with computational cost which is asymptotically, as $L \to \infty$, bounded by

$$\mathrm{O}((1+L)^{\mu} N_L^{1+\frac{2}{d}}) , \quad d = 1, 2, 3 , \tag{4.15}$$

where μ is taken from Assumption 4.6*.*

Proof. The assertions follow from the discretization error estimate in Theorem 4.5 and the basic relations (4.14). ▯

4.4. MLMC-FEM. Instead of approximating $E(u)$ directly, the MLMC-FEM is based on suitable approximations of increments on the levels $l = 1, \ldots, L$ of the hierarchy (4.1).

With the notation $u_0 := 0$, we may write

$$u_L = \sum_{l=1}^{L} (u_l - u_{l-1}) .$$

The linearity of the expectation operator $\mathbb{E}[\cdot]$ yields

$$\mathcal{K}_L \ni \mathbb{E}[u_L] = \mathbb{E}\left[\sum_{l=1}^{L} (u_l - u_{l-1})\right] = \sum_{l=1}^{L} \mathbb{E}[u_l - u_{l-1}] . \tag{4.16}$$

In the MLMC-FE method, we estimate $\mathbb{E}[u_l - u_{l-1}]$ by a level dependent number M_l of samples; i.e. we estimate $\mathbb{E}[u_L]$ by the MLMC estimator

$$E^L[u_L] := \sum_{l=1}^{L} E_{M_l}[u_l - u_{l-1}] . \tag{4.17}$$

Note that $E^L[u_L]$ differs from the MC-FE approximation $E_M[u_L]$, once different numbers of samples M_l are chosen on different levels $l = 1, \ldots, L$.

Remark 4.8. We emphasize that, in contrast to the classical MC-FEM, the MLMC-FEM is nonconforming, i.e., $E^L[u_L] \notin \mathcal{K}_L$, in general. The reason is that

$$E^L[u_L] = E_{M_L}[u_L] + \sum_{l=1}^{L-1} \left(E_{M_l}[u_l] - E_{M_{l+1}}[u_l] \right) ,$$

with $E_{M_L}[u_L] \in \mathcal{K}_L$, but $E_{M_l}[u_l] - E_{M_{l+1}}[u_l] \neq 0$, in general.

The convergence of the MLMC-FEM is guaranteed by the following result.

THEOREM 4.9. *Let Assumptions* 3.1 *and* 3.2 *hold. Then the MLMC-FE approximation $E^L[u_L]$ defined in* (4.17) *of the expectation $\mathbb{E}[u]$ of the solution $u \in L^2(\Omega; W)$ to the weak formulation* (3.14) *of the sEVI* (1.4) *admits the error bound*

$$\|\mathbb{E}[u] - E^L[u_L]\|_{L^2(\Omega;V)} \leq C(a, f, \chi) \left(h_L + \sum_{l=1}^{L} h_l M_l^{-1/2} \right) . \tag{4.18}$$

Proof. We rewrite the error to be estimated as in the proof of Theorem 4.5

according to

$$\begin{aligned}\|\mathbb{E}[u] - E^L[u_L]\|_{L^2(\Omega;V)} &= \|\mathbb{E}[u] - \mathbb{E}[u_L] + \mathbb{E}[u_L] - \sum_{l=1}^{L} E_{M_l}[u_l - u_{l-1}]\|_{L^2(\Omega;V)} \\ &\le \|\mathbb{E}[u] - \mathbb{E}[u_L]\|_{L^2(\Omega;V)} + \left\|\sum_{l=1}^{L} (\mathbb{E}[u_l - u_{l-1}] - E_{M_l}[u_l - u_{l-1}])\right\|_{L^2(\Omega;V)} \\ &=: I + II.\end{aligned}$$

We calculate the error bounds for the terms I and II separately.

Term I: By Jensen's inequality, the Cauchy–Schwarz inequality, and Proposition 4.2, we obtain

$$I \le \left(\mathbb{E}[\|u - u_L\|_V^2]\right)^{1/2} = \|u - u_L\|_{L^2(\Omega;V)} \le C(a, f, \chi)\, h_L,$$

providing the asserted bound for Term I.

Term II: By the triangle inequality, we must consider for each $l = 1, \dots, L$ the term

$$\|\mathbb{E}[u_l - u_{l-1}] - E_{M_l}[u_l - u_{l-1}]\|_{L^2(\Omega;V)}\ .$$

Each of these terms is estimated as follows:

$$\begin{aligned}\|\mathbb{E}[u_l - u_{l-1}] - E_{M_l}[u_l - u_{l-1}]\|_{L^2(\Omega;V)} &= \|(\mathbb{E} - E_{M_l})[u_l - u_{l-1}]\|_{L^2(\Omega;V)} \\ &\le M_l^{-1/2}\|u_l - u_{l-1}\|_{L^2(\Omega;V)} \\ &\le M_l^{-1/2}\left(\|u - u_l\|_{L^2(\Omega;V)} + \|u - u_{l-1}\|_{L^2(\Omega;V)}\right) \\ &\le C(a, f, \chi) M_l^{-1/2}(h_l + h_{l-1}) \\ &= 3C(a, f, \chi)\, h_l M_l^{-1/2}\ .\end{aligned}$$

Here we used Lemma 4.3, Proposition 4.2, and $h_l = 2^{-l}h_0$. Summation of these estimates for $l = 1, \dots, L$ completes the proof. ∎

The preceding result gives an error bound for the MLMC-FE approximation for *any* distribution $\{M_l\}_{l=1}^L$ of samples over the mesh levels. As in the single-level MC approximation the key question to be answered by our error analysis is the relation of meshwidth versus sample size, in order to retain the asymptotic rate of convergence $\mathrm{O}(h_l)$ from the deterministic case while minimizing the overall work.

THEOREM 4.10. *Let Assumptions* 3.1*,* 3.2*, and* 4.6 *hold. Then the inexact MLMC-FE approximation*

$$E^L[\tilde{u}_L] = \sum_{l=1}^{L} E_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}] \tag{4.19}$$

of the expectation $\mathbb{E}[u]$ *of the solution to the weak formulation* (3.14) *of the sEVI* (1.4) *with the number M_l of MC samples on mesh refinement level l given by*

$$M_l = \mathrm{O}(l^{2+2\varepsilon} 2^{2(L-l)})\,, \qquad l = 1, 2, \dots, L\,, \tag{4.20}$$

with some fixed $\varepsilon > 0$ admits the error bound

$$\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|_{L^2(\Omega;V)} \le Ch_L\|f\|_{L^2(\Omega;L^2(D))} \tag{4.21}$$

and can be evaluated at computational cost which is asymptotically, as $L \to \infty$, *bounded by*

$$\text{(4.22)} \qquad \mathrm{O}((L+1)^{\alpha+\mu+2\varepsilon} N_L^{\beta}), \quad \textit{with} \quad \begin{cases} \alpha = 2\,,\ \beta = 2\,, & d = 1\,, \\ \alpha = 3\,,\ \beta = 1\,, & d = 2\,, \\ \alpha = 2\,,\ \beta = 1\,, & d = 3\,, \end{cases}$$

where the constant implied in $\mathrm{O}(\cdot)$ *depends on* ε *and where* μ *is taken from Assumption* 4.6.

Proof. The convergence result in Theorem 4.9 suggests that we choose M_l such that the overall rate of convergence is $\mathrm{O}(h_L)$. With the choice

$$\text{(4.23)} \qquad M_l = \mathrm{O}(l^{2+2\varepsilon}(h_l/h_L)^2) = \mathrm{O}(l^{2+2\varepsilon}2^{2(L-l)}), \qquad l = 1, \ldots, L\,,$$

for some $\varepsilon > 0$, we obtain from (4.18) the asserted error bound, since for $\varepsilon > 0$ this implies

$$\begin{aligned} \sum_{l=1}^{L} h_l M_l^{-1/2} &\le C \sum_{l=1}^{L} 2^{-l} l^{-(1+\varepsilon)} 2^{(l-L)} h_0 \le C 2^{-L} h_0 \sum_{l=1}^{L} l^{-(1+\varepsilon)} \\ &\le C h_L \sum_{l=1}^{L} l^{-(1+\varepsilon)} = C(\varepsilon) h_L\,. \end{aligned}$$

It remains to estimate the computational cost. Utilizing Assumption 4.6 and (4.20), the cost is $\mathrm{O}((1+l)^{\mu} N_l M_l)$ on each level l. There results the following upper bound for the overall computational cost:

$$\sum_{l=1}^{L} \mathrm{O}((1+l)^{\mu} N_l M_l) \le C \sum_{l=1}^{L} (1+l)^{\mu} 2^{dl} l^{2+2\varepsilon} 2^{2(L-l)} \le C(1+L)^{2+\mu+2\varepsilon} N_L \sum_{l=1}^{L} 2^{(d-2)(l-L)}\,.$$

This proves the assertion. ☐

Note that for $d = 2$, 3 space dimensions the upper bounds for the computational cost of the MLMC-FEM stated in Theorem 4.10 differ from the deterministic case only by additional logarithmic terms. For all $d = 1$, 2, 3 the upper bounds of the MLMC-FEM are significantly lower than the corresponding bounds for the MC-FEM stated in Theorem 4.7. Our computations to be reported in section 5 will provide numerical evidence that these bounds are sharp and that the asymptotic superiority of the MLMC-FEM predicted by the foregoing analysis is indeed realized even on coarse meshes.

Remark 4.11. According to Theorem 4.9, the error bound (4.21) is provided by any choice of numbers of samples $(M_1, \ldots, M_L) \in \mathcal{M}_L$ with $\mathcal{M}_L \subset \mathbb{N}^L$ consisting of all integer vectors $\underline{M} = (M_1, \ldots, M_L)$ with the properties

$$\text{(4.24)} \qquad c_0 \sum_{l=1}^{L} h_l M_l^{-1/2} \le h_L\,, \qquad c_1 M_1 \ge h_L^{-2}\,, \qquad c_2 M_l \ge 1\,,\ l = 1, \ldots, L\,,$$

with arbitrary, fixed constants c_0, c_1, $c_2 > 0$ independent of L. In light of the proof of Theorem 4.10, we also recover the upper bounds (4.22) for the computational cost by selecting

$$\text{(4.25)} \qquad (M_1, \ldots, M_L) = \arg \min_{\underline{M} \in \mathcal{M}_L} \sum_{l=1}^{L} N_l M_l\,,$$

because $\underline{M} \in \mathcal{M}_L$ holds for $M_l = \lfloor C(l^{2+2\varepsilon}2^{2(L-l)})\rfloor$, $l = 1, \dots, L$ (Gaussian brackets), with arbitrary fixed $\varepsilon > 0$ and sufficiently large $C > 0$ depending on ε, c_0, c_1, c_2, and h_0 but not on l or L. Hence, selecting an (approximate) solution of the integer programming problem (4.25) with constraints (4.24) instead of $M_l = \lfloor C(l^{2+2\varepsilon}2^{2(L-l)})\rfloor$, $l = 1, \dots, L$, with some fixed $C, \varepsilon > 0$ will preserve the qualitative asymptotic bounds stated in Theorem 4.10 but might provide a quantitative reduction of computational cost. We will use this strategy in our numerical computations to be reported below.

4.5. Algebraic solution. We now discuss the evaluation of approximations $\tilde{u}_l(\omega)$ of finite element solutions $u_l(\omega)$ of deterministic obstacle problems of the form (4.4) by iterative solvers.

Assumption 4.12. There is an iterative scheme $M_l : V_l \to V_l$ for the approximate solution of deterministic obstacle problems of the form (4.4) with symmetric bilinear form $b(\omega; \cdot, \cdot)$, $\omega \in \Omega$, such that $M_l v$ can be evaluated with optimal computational cost $\mathrm{O}(N_l)$ and such that

$$J(M_l v) - J(u_l(\omega)) \le \big(1 - c_0(1 + l)^{-(\mu-1)}\big)\big(J(v) - J(u_l(\omega))\big) \tag{4.26}$$

holds for all $v \in V_l$ with some constants $c_0 > 0$ and $\mu \ge 1$, independent of $v \in V_l$, of ω, and of $l = 0, 1, \dots$.

Remark 4.13. Assumption 4.12 is fulfilled by various multigrid methods. Tai [41] proved logarithmic upper bounds of the form (4.26) with $\mu = 3$ for a class of subset decomposition methods. Badea [3] showed (4.26) with $\mu = 6$ for a projected multilevel relaxation scheme [24, section 5.1]. Badea [4] recently extended these results to obtain $\mu = 5$ for STDMMG [24, section 5.2], which, of these three multigrid methods, is the most efficient one. All these results for deterministic problems are restricted to $d = 2$ space dimensions but do require only the (minimal) H^1-regularity of the exact solution.

PROPOSITION 4.14. *Assume that an initial approximation $\tilde{u}_0(\omega) \in V_0$ with the property $\|u_0(\omega) - \tilde{u}_0(\omega)\|_V \le C_0 h_0$ with some constant $C_0 > 0$ (independent of ω but possibly depending on the data) is given and that Assumption* 4.12 *holds true. Then Assumption* 4.6 *is satisfied.*

Proof. Exploiting that the finite element spaces are nested, (4.1), we inductively construct a sequence of approximations $\tilde{u}_i(\omega) \in V_i$, $i = 0, \dots, l$. To this end, starting with the given $\tilde{u}_0 \in V_0$, we determine $\tilde{u}_i(\omega) \in V_i$ from the given $\tilde{u}_{i-1}(\omega) \in V_{i-1}$ on the previous level $i - 1$ as follows. If for some constant $C_0 > 0$ independent of ω there holds that

$$\|u_i(\omega) - \tilde{u}_{i-1}(\omega)\|_V \le C_0 2^{-i} \ , \tag{4.27}$$

then we simply set $\tilde{u}_i = \tilde{u}_{i-1}$. Otherwise, the approximation $\tilde{u}_i(\omega) = M_i^{k_i}\tilde{u}_{i-1}(\omega)$ is computed by k_i applications of the iterative solver M_i to $\tilde{u}_{i-1}(\omega)$, where k_i is chosen such that the stopping criterion

$$\|u_i(\omega) - M_i^{k_i}\tilde{u}_{i-1}(\omega)\|_V \le \tfrac{\sigma}{2}\|u_i(\omega) - \tilde{u}_{i-1}(\omega)\|_V \tag{4.28}$$

holds with some fixed $\sigma < 1$. Such a k_i exists, because M_i is convergent by (4.26) in Assumption 4.12. This process is referred to as *nested iteration* (see, e.g., [27, Chapter 5]) or *full multigrid*. In the case that (4.27) holds, we obviously have

$$\|u_i(\omega) - \tilde{u}_i(\omega)\|_V \le C_0 2^{-i} \ . \tag{4.29}$$

We assume without loss of generality that $i_0 = 0$ is the largest level $i_0 \le l$ such that (4.29) holds true, which means that (4.27) does not occur.

Utilizing (4.28), we compute

$$\begin{aligned}\|u(\omega) - \tilde{u}_l(\omega)\|_V &\le \|u(\omega) - u_l(\omega)\|_V + \|u_l(\omega) - \tilde{u}_l(\omega)\|_V \\ &\le \|u(\omega) - u_l(\omega)\|_V + \tfrac{\sigma}{2}\|u_l(\omega) - \tilde{u}_{l-1}(\omega)\|_V \\ &\le (1 + \tfrac{\sigma}{2})\|u(\omega) - u_l(\omega)\|_V + \tfrac{\sigma}{2}\|u(\omega) - \tilde{u}_{l-1}(\omega)\|_V \ .\end{aligned}$$

Successive application of this estimate for $l - i$, $i = 1, \dots, l$, instead of l then leads to

$$\|u(\omega) - \tilde{u}_l(\omega)\|_V \le (\tfrac{\sigma}{2})^l \|u_0(\omega) - \tilde{u}_0(\omega)\|_V + (1 + \tfrac{\sigma}{2}) \sum_{i=0}^{l-1} (\tfrac{\sigma}{2})^i \|u(\omega) - u_{l-i}(\omega)\|_V \ .$$

Exploiting $\|u(\omega) - \tilde{u}_0(\omega)\|_V \le C_0 h_0$, $h_l = 2^{i-l} h_i$ for $0 \le i \le l$, and the discretization error estimate (4.7), we obtain the norm error estimate in Assumption 4.6.

We now estimate the computational cost. No computations are needed in the (nongeneric) case (4.27). Hence, we assume

$$\|u_i(\omega) - \tilde{u}_{i-1}(\omega)\|_V > C_0 2^{-i} \ . \tag{4.30}$$

Utilizing (2.6) and the equivalence of the energy norm $\|\cdot\|_b = b(\omega; \cdot, \cdot)^{1/2}$ and the canonical norm $\|\cdot\|_V$ together with the upper bound (4.26) for the convergence rate of the energy error, and the upper bound (4.7) for the discretization error, it turns out that the stopping criterion (4.28) is satisfied if k_i is chosen such that

$$c\big(1 - c_0(1+i)^{-(\mu-1)}\big)^{k_i} \le \tfrac{\sigma^2}{4} \|u_i(\omega) - \tilde{u}_i(\omega)\|_V$$

holds with a suitable positive constant c which is independent of i and of μ. In light of (4.30), it is sufficient to choose k_i according to

$$k_i \ge (\log(2)i - \log(C_0/c))/ - \log(1 - c_0(1+i)^{-(\mu-1)}) \ge C(1+i)^\mu \ .$$

Hence, the computational cost on each level i is bounded by $\mathrm{O}((1+i)^\mu N_i)$. Utilizing $N_i = \mathrm{O}(2^{i-l} N_l)$, an upper bound for the overall computational cost is given by

$$\sum_{i=1}^{l} \mathrm{O}((1+i)^\mu N_i) = \mathrm{O}((1+l)^\mu N_l) \, . \qquad \square \tag{4.31}$$

Note that the additional power of $1 + l$ is caused by the mismatch between the lower and upper bounds in (2.6).

The actual computation of an approximate solution $\tilde{u}_l(\omega)$ satisfying the accuracy requirement of Assumption 4.6 can be carried out according to the proof of Proposition 4.14 (nested iteration). As the initial grid is intentionally coarse, suitable initial approximations $\tilde{u}_0(\omega)$ of $u_0(\omega)$ and thus of $u(\omega)$ can often be obtained by methods for complementarity problems with moderate size. In section 5 below, we will provide computable a posteriori bounds for the unknown algebraic error occurring in the stopping criterion (4.28).

The main result of this section is an immediate consequence of Theorems 4.7 and 4.10 together with Remark 4.13.

COROLLARY 4.15. *Let Assumptions* 3.1 *and* 3.2 *hold, assume that* $d = 2$, *and let STDMMG be used for the approximate solution of the discrete pathwise obstacle problems of the form* (4.4). *Then the resulting MC-MMG-FE approximation* $E_M[\tilde{u}_L]$ *and the resulting MLMC-MMG-FE approximation* $E^L[\tilde{u}_L]$ *of the expectation* $\mathbb{E}[u]$ *both have optimal accuracy* $\mathrm{O}(h_L)$ *in* $L^2(\Omega; V)$ *and require the computational costs* $\mathrm{O}((1+L)^5 N_L^2)$ *and* $\mathrm{O}((1+L)^{8+2\varepsilon} N_L)$, *respectively.*

Hence, utilizing the recent convergence results by Badea [4] for STDMMG in $d = 2$ space dimensions, the order of computational cost for the MLMC-FEM for the approximation of the statistical mean (or ensemble average) $\mathbb{E}[u]$ (and also for spatial correlation functions; see, e.g. [34]) turns out to be asymptotically the same as for the multigrid solution of *a single instance of the deterministic problem, on the finest mesh at refinement level* L, up to logarithmic terms. Numerical experiments indicate even mesh-independent convergence rates for STDMMG and for the recently introduced *truncated nonsmooth Newton multigrid* (TNNMG) [23, 24] as applied in the framework of nested iteration (see, e.g., Gräser and Kornhuber [24]). However, to the best of our knowledge, mathematical justification of the observed performance of TNNMG is still open.

5. Numerical experiments. Our computations are intended to provide numerical evidence that the asymptotic upper bounds for the computational cost of the MC-FEM (see section 4.3) and the MLMC-FEM (see section 4.4) as shown in Theorems 4.7 and 4.10, respectively, are sharp and that the superiority of these upper bounds for the MLMC-FEM to the MC-FEM is realized for reasonable mesh size h_L, i.e., that the implied constants in the error bounds are not excessively large.

To this end, we consider the stochastic obstacle problem (1.4) on the spatial domain $D = (-1, 1)^d$, $d = 1, 2$, with "flat" obstacle $\chi \equiv 0$, with parametric, stochastic diffusion coefficient

$$a(x, \omega) = 1 + \frac{\cos|x|^2}{10} Y_1(\omega) + \frac{\sin|x|^2}{10} Y_2(\omega) \ ,$$

and with the stochastic source term f given by

$$f(x,\omega) = \begin{cases} -8e^{2(Y_1(\omega)+Y_2(\omega))} \left(\frac{d}{2} a(x,\omega) \cdot ((4-d)|x|^2 - r^2)\right. & \\ \quad \left. + (|x|^2 - r^2)|x|^2 \left(-\frac{\sin|x|^2}{10} Y_1(\omega) + \frac{\cos|x|^2}{10} Y_2(\omega)\right)\right) \ , & |x| > r \ , \\ 4r^2 e^{2(Y_1(\omega)+Y_2(\omega))} \left(d \cdot a(x,\omega) \cdot (-1 - r^2 + |x|^2)\right. & \\ \quad \left. + (-2 - 2r^2 + |x|^2)|x|^2 \left(-\frac{\sin|x|^2}{10} Y_1(\omega) + \frac{\cos|x|^2}{10} Y_2(\omega)\right)\right) , & |x| \le r \ , \end{cases}$$

denoting

$$r = r(Y_1(\omega), Y_2(\omega)) := 0.7 + \frac{Y_1(\omega) + Y_2(\omega)}{10} \ ,$$

and uniformly distributed random variables $Y_1, Y_2 \sim \mathcal{U}(-1, 1)$. Then, for given $\omega \in \Omega$, the exact solution of the resulting pathwise problem is given by

$$u(x, \omega) = \max\{\left(|x|^2 - r^2\right) e^{Y_1(\omega)+Y_2(\omega)}, 0\}^2 \ , \qquad x \in D \ .$$

Each pathwise problem is discretized by finite elements as described in section 4.1. To build up the hierarchy (4.1), we start from the coarse partition $\mathcal{T}_0$ consisting of four intervals with length $h_0 = 1/4$ for $d = 1$ and $\mathcal{T}_0$ resulting from uniform refinement of a partition of D into two congruent triangles for $d = 2$ space dimensions. Approximate solution of the resulting discrete pathwise problems of the form (4.4) up to the norm error estimate in Assumption 4.6 is performed by a V-cycle with three pre- and three postsmoothing steps of the TNNMG method [23, 24] with nested iteration (see [27, Chapter 5] and the proof of Proposition 4.14). The reason is that TNNMG is easier to implement and usually converges faster than STDMMG [24]. Denoting one step of TNNMG on refinement level i by M_i, the stopping criterion (4.28) is replaced by the verifiable condition

$$\|M_i^{k_i+1}\tilde{u}_{i-1}(\omega) - M_i^{k_i}\tilde{u}_{i-1}(\omega)\|_V \leq s\tfrac{\sigma}{2}\|M_i\tilde{u}_{i-1}(\omega) - \tilde{u}_{i-1}(\omega)\|_V$$

with a safety factor $s \leq (1-q)/(1+q)$. This condition relies on the a posteriori error estimate

$$(1-q)\|u_i(\omega) - v\|_V \leq \|M_i v - v\|_V \leq (1+q)\|u_i(\omega) - v\|_V \;, \quad v \in V_i \;,$$

involving the (unknown) convergence rate $q < 1$ of M_i. We use $s = 0.1$ and $\sigma = \frac{1}{2}$ in our computations.

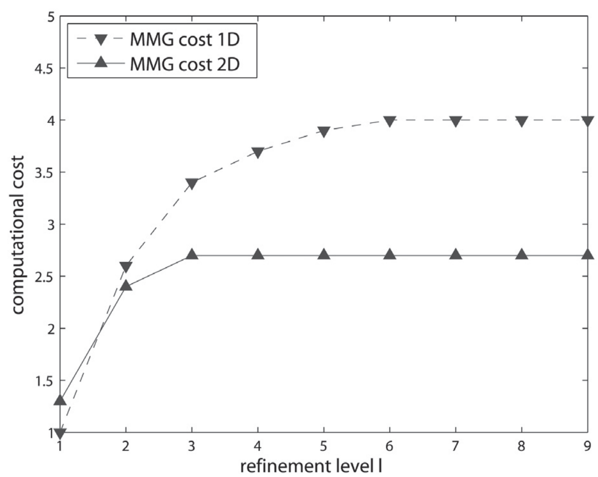

FIG. 1. *Averaged computational cost per unknown for the TNNMG solution of obstacle problems as occurring in an MLMC-FE computation over the number of refinement levels l.*

To illustrate the performance of the resulting algebraic solver, the cost for one step of TNNMG on level i is set to N_i (ignoring constants) so that for each fixed realization ω^l the computational cost on level l is $\sum_{i=1}^{l} k_i N_i$ with the number of steps $k_i = k_i(\omega^l)$ possibly depending on the actual realization ω^l. Figure 1 shows the *average computational cost per unknown* on the levels $l = 1, \ldots, 9$. More precisely, it shows the computational cost averaged over all samples ω^l as occurring in MC computations on the level $l = 1, \ldots, 9$ in the course of an MLMC-FE computation in $d = 1$ and $d = 2$ space dimensions on the final level $L = 9$ (see below) divided by the number of unknowns N_l. According to Figure 1, the average computational cost per unknown saturates for increasing $l = 1, \ldots, 9$, indicating that the overall cost on level l is bounded by about $4.0N_l$ and $2.7N_l$ uniformly for all $l = 1, \ldots$ in $d = 1$ and $d = 2$

Table 1
Number of samples $M_1, \dots, M_L$ for $L = 2, \dots, 9$ in $d = 1$ space dimension.

L	2	3	4	5	6	7	8	9
M_1	29	252	1601	8692	43025	200820	900780	3928500
M_2	23	208	1322	7175	35517	165780	743580	3242900
M_3		83	525	2848	14095	65790	295090	1287000
M_4			209	1130	5594	26110	117110	510700
M_5				449	2220	10360	46470	202700
M_6					881	4110	18440	80400
M_7						1630	7320	31900
M_8							2900	12700
M_9								5000

Table 2
Number of samples $M_1, \dots, M_L$ for $L = 2, \dots, 7$ in $d = 2$ space dimensions.

L	2	3	4	5	6	7
M_1	38	427	3355	22238	133090	744340
M_2	20	222	1745	11564	69210	387070
M_3		56	437	2891	17300	96770
M_4			109	723	4330	24190
M_5				181	1080	6050
M_6					270	1510
M_7						380

space dimensions, respectively, no matter how the actual realization ω^l is chosen. This means that, numerically, Assumption 4.6 holds with the optimal parameter $\mu = 0$. We obtained essentially the same results for STDMMG.

The number of samples $M_1, \dots, M_L$ to be used in the MLMC-FEM is selected according to Remark 4.11. More precisely, we select the approximate solution of the integer programming problem (4.25) with constraints (4.24) and parameters $c_0 = 3$, $c_1 = c_2 = 1$ as obtained by rounding up the solution of the corresponding continuous constrained minimization problem which in turn is computed numerically by a MATLAB routine. The resulting numbers of samples $M_1, \dots, M_L$ are reported in Tables 1 and 2 for the above model problem in $d = 1$ and $d = 2$ space dimensions, respectively. Once $M_1, \dots, M_L$ have been determined, we choose the number of samples $M = M_1$ for the single-level MC-FEM. Recall that $M = M_1 \geq h_L^{-2}$ holds by (4.24).

The evaluation of the statistical error

$$\|\mathbb{E}[u_L] - E[\tilde{u}_L]\|_{L^2(\Omega,V)} = \left(\mathbb{E}\left[\|\mathbb{E}[u_L] - E[\tilde{u}_L]\|_V^2\right]\right)^{1/2} \tag{5.1}$$

associated with the discrete approximations $E[\tilde{u}_L] = E_M[\tilde{u}_L]$, $E^L[\tilde{u}_L]$ requires the approximate evaluation of the expectation of the random variable $\|\mathbb{E}[u_L] - E[\tilde{u}_L]\|_V^2$. We approximate $\mathbb{E}\left[\|\mathbb{E}[u_L] - E[\tilde{u}_L]\|_V^2\right]$ by an MC method or, more precisely, by the sample average of 10 numerically computed, independent realizations of $\|\mathbb{E}[u_L] - E[\tilde{u}_L]\|_V^2$.

We now report on the numerical solution of the model problem introduced above in $d = 1$ space dimension. Figure 2 nicely illustrates the $\mathrm{O}(h_L)$ behavior of the statistical errors of the MC-FEM (-•–•-) and the MLMC-FEM (-♦–♦-) indicated by the dashed line. Note that the MC-FEM is slightly more accurate than the MLMC-FEM by a factor ranging from 1.05 to 2.04 over the levels $L = 2, \dots, 9$. The corresponding

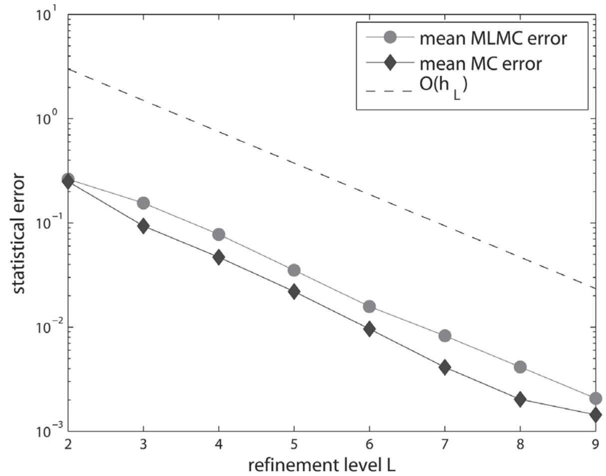

FIG. 2. *Statistical error* (5.1) *of the MC-FEM and the MLMC-FEM over the number of refinement levels L in $d = 1$ space dimension.*

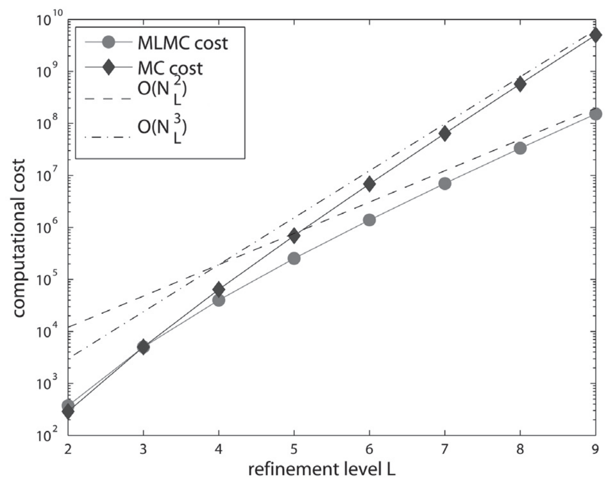

FIG. 3. *Computational cost of the MC-FEM and the MLMC-FEM over the number of refinement levels L in $d = 1$ space dimension.*

computational cost $N_L M$ of the MC-FEM (-•–•-) and $\sum_{l=1}^{L} N_l M_l$ of the MLMC-FEM (-◆–◆-) over the refinement levels $L = 2, \ldots, 9$ is depicted in Figure 3. It turns out that the cost of the MC-FEM asymptotically behaves like $\mathrm{O}(N_L^3)$ (dotted line), while the MLMC-FEM requires only $\mathrm{O}(N_L^2)$ point operations (dashed line). For this moderate number of refinement levels, the logarithmic term occurring in the theoretical upper bound (4.22) for $\mu = 0$ is not visible. These computational results indicate that the asymptotic upper bounds for the cost stated in Theorems 4.7 and 4.10 are sharp. Moreover, the predicted asymptotic (as $L \to \infty$) superiority of these asymptotic bounds for the MLMC-FEM is realized already for rather coarse meshes in these

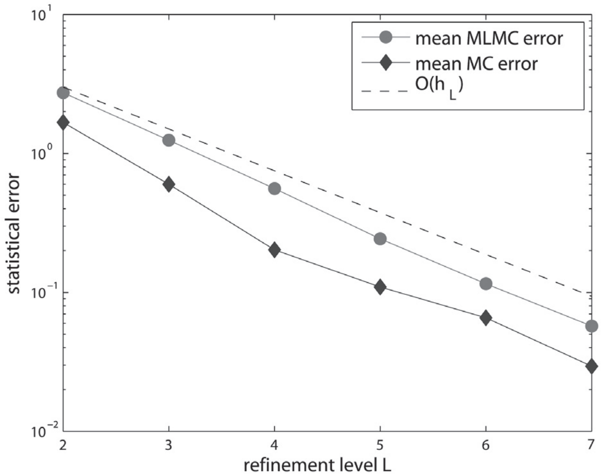

FIG. 4. *Statistical error* (5.1) *of the MC-FEM and the MLMC-FEM over the number of refinement levels L in $d = 2$ space dimensions.*

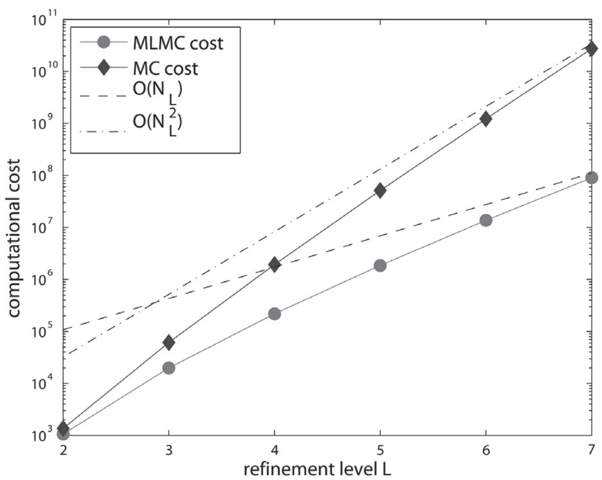

FIG. 5. *Computational cost of the MC-FEM and the MLMC-FEM over the number of refinement levels L in $d = 2$ space dimensions.*

examples.

The corresponding results for $d = 2$ space dimensions are shown in Figures 4 and 5. By Figure 4, the statistical error again behaves like $\mathrm{O}(h_L)$ (dashed line) and the MC-FEM (-•–•-) is slightly more accurate than the MLMC-FEM (-♦–♦-) by a factor ranging from 1.38 to 2.48 over the levels $L = 2, \dots, 7$. While in $d = 1$ space dimension the evolution of computational costs with increasing L fully entered the asymptotic regime, this is not the case for $d = 2$, because the maximal number of refinement steps is now limited to $L = 7$ (instead of $L = 9$), due to limited resources in memory and computing time (mainly for the MC-FEM!). Nevertheless, the results depicted in Figure 5 indicate that the computational cost of the MC-FEM (-•–•-) is of order $O(N_L^2)$ and that the MLMC-FEM (-⧫–⧫-) provides approximations with $O(h_l)$ accuracy at optimal computational cost $O(N_L)$. This gives some computational evidence that the asymptotic upper bounds for the computational cost stated in Theorems 4.7 and 4.10 are sharp. Again, asymptotic superiority of the MLMC-FEM to the MC-FEM is already realized for coarse meshes.

Acknowledgment. The authors thank Carsten Gräser (FU Berlin) for fruitful discussions. In particular, he suggested computing the number of samples M_l according to Remark 4.11.

REFERENCES

[1] R. J. Adler, *The Geometry of Random Fields*, John Wiley & Sons, Chichester, UK, 1981.

[2] H. W. Alt, L. A. Caffarelli, and A. Friedman, *Variational problems with two phases and their free boundaries*, Trans. Amer. Math. Soc., 282 (1984), pp. 431–461.

[3] L. Badea, *Convergence rate of a Schwarz multilevel method for the constrained minimization of nonquadratic functionals*, SIAM J. Numer. Anal., 44 (2006), pp. 449–477.

[4] L. Badea, *Global convergence rate of a standard multigrid method for variational inequalities*, IMA J. Numer. Anal., 34 (2014), pp. 197–216.

[5] A. Barth, *A finite element method for martingale-driven stochastic partial differential equations*, Comm. Stoch. Anal., 4 (2010), pp. 355–375.

[6] A. Barth and A. Lang, *Simulation of stochastic partial differential equations using finite element methods*, Stochastics, 84 (2012), pp. 217–231.

[7] A. Barth and A. Lang, *L^p and almost sure convergence of a Milstein scheme for stochastic partial differential equations*, Stochastic Process. Appl., 123 (2013), pp. 1563–1587.

[8] A. Barth, C. Schwab, and N. Zollinger, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.

[9] J. Bey, *Finite-Volumen- und Mehrgitterverfahren für elliptische Randwertprobleme*, Teubner, Stuttgart, 1998.

[10] C. Bierig and A. Chernov, *Analysis of Multilevel Variance Estimators in Multilevel Monte Carlo Methods and Application for Random Obstacle Problems*, Technical Report 1309, INS Bonn, Bonn, Germany, 2013.

[11] F. A. Bornemann, B. Erdmann, and R. Kornhuber, *A posteriori error estimates for elliptic problems in two and three space dimensions*, SIAM J. Numer. Anal., 33 (1996), pp. 1188–1204.

[12] D. Braess, *Finite Elements*, 3rd ed., Springer, Berlin, 2002.

[13] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 2nd ed., Springer, Berlin, 2002.

[14] J. Charrier, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 216–246.

[15] J. Charrier, R. Scheichl, and A. L. Teckentrup, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, SIAM J. Numer. Anal., 51 (2013), pp. 322–352.

[16] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North–Holland, Amsterdam, 1978.

[17] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia Math. Appl. 44, Cambridge University Press, Cambridge, UK, 1992.

[18] R. S. Falk, *Error estimates for the approximation of a class of variational inequalities*, Math. Comp., 28 (1974), pp. 963–971.

[19] R. Forster and R. Kornhuber, *A polynomial chaos approach to stochastic variational inequalities*, J. Numer. Math., 18 (2010), pp. 235–255.

[20] M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT, 49 (2009), pp. 343–356.

[21] M. Giles, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, in Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer, Berlin, 2008, pp. 343–358.

[22] M. B. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.

[23] C. Gräser, *Convex Minimization and Phase Field Models*, Ph.D. thesis, Freie Universität Berlin, Berlin, Germany, 2011.

[24] C. Gräser and R. Kornhuber, *Multigrid methods for obstacle problems*, J. Comput. Math., 27 (2009), pp. 1–44.
[25] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monogr. Stud. Math. 24, Pitman, Boston, MA, 1985.
[26] J. Gwinner, *A class of random variational inequalities and simple random unilateral boundary value problems—existence, discretization, finite element approximation*, Stoch. Anal. Appl., 18 (2000), pp. 967–993.
[27] W. Hackbusch, *Multigrid Methods and Applications*, Springer, Berlin, 1985.
[28] S. Heinrich, *Multilevel Monte Carlo methods*, in Large-Scale Scientific Computing, Lecture Notes in Comput. Sci. 2179, S. Margenov, J. Wasniewski, and P. Y. Yalamov, eds., Springer, Berlin, 2001, pp. 58–67.
[29] D. Kinderlehrer and G. Stampacchia, *Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, 1980.
[30] R. Kornhuber, *Monotone multigrid methods for elliptic variational inequalities* I, Numer. Math., 69 (1994), pp. 167–184.
[31] M. Kovács, S. Larsson, and F. Saedpanah, *Finite element approximation of the linear stochastic wave equation with additive noise*, SIAM J. Numer. Anal., 48 (2010), pp. 408–427.
[32] J. Mandel, *A multilevel iterative method for symmetric, positive definite linear complementarity problems*, Appl. Math. Optim., 11 (1984), pp. 77–95.
[33] B. Matérn, *Spatial Variation*, 2nd ed., Lecture Notes in Statist. 36, Springer-Verlag, Berlin, 1986.
[34] S. Mishra and C. Schwab, *Sparse tensor multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data*, Math. Comp., 81 (2012), pp. 1979–2018.
[35] R. L. Naff, D. F. Haley, and E. Sudicky, *High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media* 1. *Methodology and flow results*, Water Resour. Res., 34 (1998), pp. 663–677.
[36] R. L. Naff, D. F. Haley, and E. Sudicky, *High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media* 2. *Transport results*, Water Resour. Res., 34 (1998), pp. 679–697.
[37] J.-F. Rodriguez, *Obstacle Problems in Mathematical Physics*, Elsevier, Amsterdam, 1987.
[38] O. Sander, C. Klapproth, J. Youett, R. Kornhuber, and P. Deuflhard, *Towards an Efficient Numerical Simulation of Complex* 3*d Knee Joint Motion*, manuscript.
[39] C. Schwab and C. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer., 20 (2011), pp. 291–467.
[40] C. Schwab and R. A. Todor, *Karhunen-Loève approximation of random fields by generalized fast multipole methods*, J. Comput. Phys., 217 (2006), pp. 100–122.
[41] X.-C. Tai, *Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities*, Numer. Math., 93 (2003), pp. 755–786.