View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Repository: Freie Universität Berlin (FU), Math Department (fu_mi_publications)

SIAM J. NUMER. ANAL.
Vol. 56, No. 4, pp. 1987–2007

© 2018 Society for Industrial and Applied Mathematics

ADAPTIVE MULTILEVEL MONTE CARLO METHODS FOR STOCHASTIC VARIATIONAL INEQUALITIES*

RALF KORNHUBER† AND EVGENIA YOUETT†

Abstract. While multilevel Monte Carlo (MLMC) methods for the numerical approximation of partial differential equations with random coefficients enjoy great popularity, combinations with spatial adaptivity seem to be rare. We present an adaptive MLMC finite element approach based on deterministic adaptive mesh refinement for the arising "pathwise" problems and outline a convergence theory in terms of desired accuracy and required computational cost. Our theoretical and heuristic reasoning together with the efficiency of our new approach are confirmed by numerical experiments.

Key words. finite elements, adaptive mesh refinement, uncertainty quantification

AMS subject classifications. 65N30, 65N50, 65N55

DOI. 10.1137/16M1104986

1. Introduction.

Uncertainty quantification is a well-established and flourishing field in numerical analysis and scientific computing that connects theoretical challenges with a multitude of practical applications. While stochastic Galerkin approaches (cf., e.g., [3, 4, 48]) turned out as methods of choice for low dimensional uncertainties, Monte Carlo (MC) type of methods prove advantageous for high dimensional, highly nonlinear problems. While the classical MC method is very robust and extremely simple, sampling of stochastic data entails the numerical solution of numerous deterministic problems, which makes performance the main weakness of this approach. A big step toward efficiency was made by Giles [28], who combined MC with multigrid techniques by introducing suitable hierarchies of subproblems associated with corresponding mesh hierarchies. Since then, multilevel Monte Carlo (MLMC) methods have become a powerful tool in a variety of applications and a flourishing field of mathematical research. We refer to elliptic problems with random coefficients [8, 17, 18, 45], random elliptic problems with multiple scales [1], parabolic random problems [7], and random elliptic variational inequalities [10, 40] and to [29] for a detailed overview.

Various approaches have been made to further enhance the efficiency of MLMC. For a given, quasi-uniform mesh hierarchy, Collier et al. [19] and Haji-Ali et al. [35] aim at reducing the computational cost of MLMC by optimizing the actual selection of meshes from this hierarchy and other MLMC parameters.

Another, in a sense complementary approach, to reducing the required computational cost of MLMC is to apply adaptive mesh refinement techniques.

Time discretization of an Itô stochastic differential equation by an self-adaptively chosen hierarchy of time steps has been suggested by Hoel et al. [36, 37] and a sim-

*Received by the editors November 23, 2016; accepted for publication (in revised form) April 16, 2018; published electronically July 3, 2018.

http://www.siam.org/journals/sinum/56-4/M110498.html

Funding: The work of the authors was supported by the German Ministry for Education and Research (BMBF) through grant *Wear Simulation and Shape Optimization of Knee Implants. Subproject* 4: *Uncertainty Quantification* and by the German Research Foundation (DFG) through grant CRC 1114 *Scaling Cascades in Complex Systems*, Project B01 *Fault Networks and Scaling Properties of Deformation Accumulation.*

†Institut für Mathematik, Freie Universität Berlin, 14195 Berlin, Germany (kornhuber@math.fu-berlin.de, babushkina@2edat.fu-berlin.de).

ilar approach was presented by Gerstner and Heinz [27], including applications in computational finance.

Less appears to be known for partial differential equations with random coefficients. While a posteriori error estimation and adaptive mesh refinement have quite a history in finite element approximation of deterministic partial differential equations (cf., e.g., [2, 47]), related adaptive concepts for MLMC methods seem to be rare. Only recently, Eigel, Merdon, and Neumann [25] suggested an algorithm for constructing an adaptively refined hierarchy of meshes based on expectations of "pathwise" local error indicators and illustrated its properties by numerical experiments. Elfverson, Hellman, and Målqvist [26] suggested a sample-adaptive MLMC method for approximate failure probability functionals and Detommaso, Dodwell, and Scheichl [21] introduced continuous level Monte Carlo (CLMC) treating the level as a continuous variable as a general framework for sample-adaptive level hierarchies.

In this paper, we follow a novel approach, utilizing a whole family of different pathwise mesh hierarchies associated with different MC samples $\omega \in \Omega$. More precisely, for a given final tolerance $Tol > 0$, we choose a sequence of tolerances $Tol_1 > \cdots > Tol_L = Tol$ and approximate each of the different pathwise deterministic problems arising for each of the different samples $\omega \in \Omega$ on each MLMC level l up to the accuracy Tol_l by finite elements on a different, adaptively refined pathwise mesh. We emphasize that any deterministic refinement strategy can be used for this purpose. The computation of sample averages is finally performed on an inductively constructed global mesh consisting of the union of simplices from all these pathwise meshes resulting from the different samples.

Based on existing results on elliptic variational inequalities [33, 38] and on general MLMC methods [18, 28], we outline an abstract convergence theory for adaptive MLMC Galerkin approximations of the expected solution in an abstract Hilbert space setting. Error estimates are formulated in terms of the desired accuracy Tol and the required computational cost. Extensions to bounded linear as well as Fréchet differentiable functionals can be obtained from corresponding results in [17, 45]. The general theory is then applied to MLMC finite element methods. In the case of uniform refinement we recover an enhanced version of existing results from [40] and we discuss the assumptions of our abstract theory in light of existing convergent adaptive algorithms for deterministic elliptic variational inequalities [15, 43] and optimality results for linear variational problems [11, 16, 41, 44]. The implementation of the resulting adaptive MLMC finite element methods is carried out in the software environment DUNE [12]. Numerical experiments illustrate our theoretical findings and the underlying heuristic reasoning. For problems with highly localized random source term, we observe a significant reduction of computational cost as compared to uniformly refined meshes. Optimal bounds for the computational cost are observed in all our numerical experiments. Theoretical justification will be the subject of future research.

The paper is organized as follows. Section 2 contains the formulation of pathwise elliptic variational inequalities together with some well-known existence and uniqueness results. In section 3 we present our abstract framework of adaptive MLMC Galerkin methods together with error estimates and upper bounds for the required computational cost. These abstract results are applied to finite element approximations in the next section 4 and numerical experiments are reported in the final section 5.

2. A random variational problem. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space with Ω denoting a sample space and let $\mathcal{A} \in 2^\Omega$ be the σ-algebra of all possible

events associated with a finite probability measure $\mathbb{P} : \mathcal{A} \to [0,1]$ on Ω. As usual, $\mathbb{E}[\xi] = \int_\Omega \xi \, d\mathbb{P}$ denotes the expectation of a random variable ξ and $L^2(\Omega)$ denotes the Hilbert space of square integrable random variables on Ω.

For a given separable Hilbert space H, equipped with the scalar product $(\cdot,\cdot)_H$ and the associated norm $\|\cdot\|_H = (\cdot,\cdot)_H^{1/2}$, we introduce the Bochner-type space $L^2(\Omega,\mathcal{A},\mathbb{P};H)$ of $\mathbb{P}$-measurable mappings $v : \Omega \to H$ with the property $\int_\Omega \|v\|_H^2 \, d\mathbb{P}(\omega) < \infty$. We will use the abbreviation $L^2(\Omega;H) = L^2(\Omega,\mathcal{A},\mathbb{P};H)$. It is easily seen that $L^2(\Omega;H)$ is also a Hilbert space with the scalar product

$$(v,w)_{L^2(\Omega;H)} = \int_\Omega (v,w)_H \, d\mathbb{P}(\omega), \qquad v,w \in L^2(\Omega;H),$$

and the associated norm $\|\cdot\|_{L^2(\Omega;H)} = (\cdot,\cdot)_{L^2(\Omega;H)}^{1/2}$. The expectation in $L^2(\Omega;H)$ is defined by

$$\mathbb{E}[v] = \int_\Omega v(\omega) \, d\mathbb{P}(\omega) \in H, \qquad v \in L^2(\Omega;H).$$

Let $a(\omega;\cdot,\cdot)$ and $\ell(\omega;\cdot)$, $\omega \in \Omega$, denote families of bilinear forms and linear functionals on H, respectively. For a given subset $K \subset H$ and any fixed realization $\omega \in \Omega$, we consider the pathwise variational inequality

$$u(\omega) \in K : \qquad a(\omega;u(\omega),v-u(\omega)) \geq \ell(\omega;v-u(\omega)) \qquad \forall v \in K. \tag{1}$$

Note that in the unconstrained case $K = H$ the inequality (1) can be equivalently rewritten as the variational equality

$$u(\omega) \in H : \qquad a(\omega;u(\omega),v) = \ell(\omega;v) \qquad \forall v \in H. \tag{2}$$

Assumption 2.1. The subset K is nonempty, closed, and convex. For each realization $\omega \in \Omega$ we have $\ell(\omega;\cdot) \in H'$ and $a(\omega;\cdot,\cdot)$ is bounded and coercive in the sense that

$$\gamma(\omega)\|v\|_H^2 \leq a(\omega;v,v), \quad a(\omega;v,w) \leq \Gamma(\omega)\|v\|_H\|w\|_H \qquad \forall v,w \in H \tag{3}$$

holds with $\gamma(\omega) \geq \gamma_0 > 0$ a.e. in Ω, and $\Gamma \in L^\infty(\Omega)$. For all fixed v, $w \in H$ the mappings $a(\cdot;v,w)$, $\ell(\cdot;v)$ are measurable and $\ell \in L^2(\Omega;H')$.

Assumption 2.1 yields the existence, uniqueness, and regularity of pathwise solutions (cf., e.g., [38, Theorem 2.1] and [33, Proposition 1.2]).

PROPOSITION 2.1. *Let Assumption* 2.1 *hold. Then the pathwise problem* (1) *admits a unique solution for each $\omega \in \Omega$, the solution map $u : \Omega \mapsto H$ is measurable with respect to the Borel σ-algebra in H, and $u \in L^2(\Omega;H)$.*

Note that $u \in L^2(\Omega;H)$ implies $\mathbb{E}[u] \in H$. It also follows that

$$u \in \mathcal{K} = \left\{ v \in L^2(\Omega;H) \mid v(\omega) \in K \text{ a.e. in } \Omega \right\} \subset L^2(\Omega;H)$$

is the unique solution of the "mean-square" variational inequality

$$u \in \mathcal{K} : \qquad \mathbb{E}[a(\cdot;u,v-u)] \geq \mathbb{E}[\ell(\cdot;v-u)] \qquad \forall v \in \mathcal{K}. \tag{4}$$

To fix the ideas, we will often concentrate on the bilinear form

$$a(\omega;v,w) = \int_D \alpha(x,\omega)\nabla v(x) \cdot \nabla w(x) \, dx$$

and the functional

$$\ell(\omega; v) = \int_D f(x, \omega) v(x)\, dx$$

on the Sobolev space $H = H_0^1(D)$ of weakly differentiable functions defined on a Lipschitz domain $D \in \mathbb{R}^d$, $d = 1, 2, 3$, and the subset

$$K = \{v \in H \mid v(x) \geq 0 \text{ a.e. in } D\}. \tag{5}$$

Note that random obstacles $\chi \in L^2(\Omega; H_0^1(D))$ can be traced back to the case (5) by introducing the new variable $w = u - \chi$. For a detailed discussion of sufficient conditions on the coefficient α and the right-hand side f for existence and uniqueness of pathwise solutions, we refer to section 4.

The remainder of this paper is devoted to the efficient approximation of the expectation $\mathbb{E}[u]$ of the family of pathwise solutions $u(\omega)$, $\omega \in \Omega$, of (1).

3. Adaptive multilevel Monte Carlo Galerkin methods. For given initial tolerance $0 < Tol_1 < 1$ and reduction factor $q < 1$ we define a sequence of tolerances by

$$Tol_l = qTol_{l-1}, \quad l = 2, \ldots, L, \tag{6}$$

with the final desired accuracy $Tol = Tol_L$. For each $\omega \in \Omega$ we choose an associated hierarchy of subspaces $S_l(\omega) \subset H$, i.e.,

$$S_1(\omega) \subset S_2(\omega) \subset \cdots \subset S_L(\omega) \subset H, \tag{7}$$

with finite dimensions $N_l(\omega)$ and nonempty, closed, convex subsets $K_l(\omega) \subset S_l(\omega)$, $l = 1, \ldots, L$. We consider the family of pathwise Galerkin approximations

$$u_l(\omega) \in K_l(\omega): \quad a(\omega; u_l(\omega), v - u_l(\omega)) \geq \ell(\omega; v - u_l(\omega)) \quad \forall v \in K_l(\omega), \qquad \omega \in \Omega. \tag{8}$$

Assumption 3.1. For all $l = 1, \ldots, L$ the set-valued map $\Omega \ni \omega \mapsto K_l(\omega) \in H$ is measurable and there is a $w_l \in L^2(\Omega; H)$ such that $w_l(\omega) \in K_l(\omega)$ holds for all $\omega \in \Omega$.

In combination with Assumption 2.1, the Assumption 3.1 yields the existence, uniqueness, and regularity of approximate pathwise solutions (cf., e.g., [34, Theorems 2.3 and 2.7]).

PROPOSITION 3.1. *Let Assumptions* 2.1 *and* 3.1 *hold. Then there is a unique solution* $u_l(\omega) \in K_l(\omega)$ *of* (8) *for each* $l = 1, \ldots, L$ *and* $\omega \in \Omega$, *the discretized solution map* $u_l : \Omega \mapsto S_l(\omega) \subset H$ *is measurable, and* $u_l \in L^2(\Omega; H)$.

Before we approximate the expectation $\mathbb{E}[u]$ in terms of (approximations of) $u_l(\omega)$, $\omega \in \Omega$, let us state some assumptions on $u_l(\omega)$ and thus implicitly on the approximating family of spaces $S_l(\omega)$.

Assumption 3.2. For all $l = 1, \ldots, L$ the family $u_l(\omega)$, $\omega \in \Omega$, satisfies the discretization error estimate

$$\|u - u_l\|_{L^2(\Omega;H)} \leq \tfrac{1}{2\sqrt{2}} Tol_l. \tag{9}$$

In general, the exact solution $u_l(\omega)$ of variational inequality (8) is not available but can be only approximated up to a certain tolerance by an iterative solver.

Assumption 3.3. For all $l = 1, \dots, L$ and each $\omega \in \Omega$, an approximate solution $\tilde{u}_l(\omega) \in S_l(\omega)$ of the pathwise problem (8) can be computed with accuracy

$$\|u_l(\omega) - \tilde{u}_l(\omega)\|_H \leq \tfrac{1}{2\sqrt{2}} Tol_l, \tag{10}$$

$\tilde{u}_l : \Omega \mapsto S_l(\omega) \subset H$ is measurable, and $\tilde{u}_l \in L^2(\Omega; H)$.

Then the expectation $\mathbb{E}(u)$ is approximated by the inexact MLMC Galerkin method

$$\mathbb{E}^L[\tilde{u}_L] = \sum_{l=1}^{L} \mathbb{E}_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}] \tag{11}$$

with $\tilde{u}_0 = 0$ and suitable $(M_l) \in \mathbb{N}^L$. On each level l, we utilize the MC approximation

$$\mathbb{E}_M[v] = \frac{1}{M} \sum_{i=1}^{M} v_i(\omega), \qquad M \in \mathbb{N}, \quad v \in L^2(\Omega; H), \tag{12}$$

of $\mathbb{E}[v]$ by independent, identically distributed copies $v_i(\omega)$ of $v(\omega)$, $i = 1, \dots, M$.

A basic representation of the error of MC methods is stated in the following lemma

LEMMA 3.1. *The MC approximation* (12) *of the expectation* $\mathbb{E}[v]$ *satisfies the identity*

$$\|\mathbb{E}[v] - \mathbb{E}_M[v]\|_{L^2(\Omega;H)} = M^{-1/2} V[v]^{1/2} \tag{13}$$

denoting

$$V[v] = \mathbb{E}[\|\mathbb{E}[v] - v\|_H^2] \leq \|v\|_{L^2(\Omega;H)}^2. \tag{14}$$

Before we present an error estimate for the inexact MLMC method, we state an extension of Lemma 3.1 taken from [10, Theorem 3.1].

LEMMA 3.2. *The inexact MLMC Galerkin approximation* $\mathbb{E}^L[\tilde{u}_L]$ *satisfies*

$$\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H)}^2 = \|\mathbb{E}[u - \tilde{u}_L]\|_H^2 + \sum_{l=1}^{L} M_l^{-1} V[\tilde{u}_l - \tilde{u}_{l-1}]. \tag{15}$$

We now prove an error bound for the inexact MLMC Galerkin method.

THEOREM 3.1. *Let Assumptions* 2.1 *and* 3.1–3.3 *hold. Then the inexact MLMC Galerkin approximation* $\mathbb{E}^L[\tilde{u}_L]$ *of the expected value* $\mathbb{E}[u]$ *satisfies the error estimate*

(16)

$$\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H)}^2$$

$$\leq 3M_1^{-1} \left(\tfrac{1}{4} Tol_1^2 + V[u]\right) + \tfrac{1}{2} \left(1 + \left(1 + q^{-1}\right)^2 \sum_{l=2}^{L} M_l^{-1} q^{2(l-L)}\right) Tol^2.$$

Proof. We estimate the terms on the right-hand side of the identity (15). First we get

$$\mathbb{E}[\|u - \tilde{u}_L\|_H] \le \|u - \tilde{u}_L\|_H \le \|u - u_L\|_H + \|u_L - \tilde{u}_L\|_H \le 2^{-1/2} Tol_L$$

utilizing the triangle inequality together with Assumptions 3.2 and 3.3. Then, for $l = 2, \dots, L$ we have

$$\begin{aligned} V[\tilde{u}_l - \tilde{u}_{l-1}] &\le \|\tilde{u}_l - \tilde{u}_{l-1}\|^2_{L^2(\Omega;H)} \\ &\le \big(\|\tilde{u}_l - u_l\|_{L^2(\Omega;H)} + \|u_l - u\|_{L^2(\Omega;H)} \\ &\qquad + \|u - u_{l-1}\|_{L^2(\Omega;H)} + \|u_{l-1} - \tilde{u}_{l-1}\|_{L^2(\Omega;H)}\big)^2 \\ &\le \tfrac{1}{2}(1 + q^{-1})^2 Tol_l^2, \end{aligned}$$

again by Assumptions 3.2 and 3.3 and (6). Finally, for $l = 1$, we obtain the estimate

$$\begin{aligned} V[\tilde{u}_1] &= V[(\tilde{u}_1 - u_1) + (u_1 - u) + u] \\ &\le 3(\|\tilde{u}_1 - u_1\|^2_{L^2(\Omega;H)} + \|u_1 - u\|^2_{L^2(\Omega;H)} + V[u]) \\ &\le 3(\tfrac{1}{4} Tol_1^2 + V[u]). \end{aligned}$$

Inserting the above estimates into (15), we obtain

$$\begin{aligned} &\big\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\big\|^2_{L^2(\Omega;H)} \\ &\qquad \le \tfrac{1}{2} Tol^2 + 3M_1^{-1}\left(\tfrac{1}{4} Tol_1^2 + V[u]\right) + \tfrac{1}{2}\left(1 + q^{-1}\right)^2 \sum_{l=2}^{L} M_l^{-1} Tol_l^2. \end{aligned}$$

As a consequence of (6), we have $Tol_l = q^{l-L} Tol$ and the assertion follows. ∎

The error estimate (16) clearly implies that the desired accuracy Tol is obtained for sufficiently large numbers of samples M_l, $l = 1, \dots, L$.

We now investigate the computational cost for the evaluation of $\mathbb{E}^L[\tilde{u}_L]$. Assuming that the evaluation of the inexact solution of the discrete pathwise problems (8) dominates overall work, the computational cost is defined by

$$\sum_{l=1}^{L} \sum_{i=1}^{M_l} cost(\tilde{u}_{l,i}(\omega)), \tag{17}$$

where $cost(\tilde{u}_{l,i}(\omega))$ stands for the computational cost of one evaluation of $\tilde{u}_{l,i}(\omega)$ measured in the number of floating-point operations. We relate $cost(\tilde{u}_{l,i}(\omega))$ to the dimension $N_{l,i}(\omega)$ of $S_{l,i}(\omega)$.

Assumption 3.4. For all $l = 1, \dots, L$ and each $\omega \in \Omega$, an approximation $\tilde{u}_l(\omega)$ of the solution $u_l(\omega)$ of (8) can be evaluated at computational cost bounded by

$$c_0(1 + \log(N_l(\omega)))^{\mu} N_l(\omega)$$

with positive constants c_0, μ independent of Tol_l, $N_l(\omega)$, and $\omega \in \Omega$.

In order to obtain a bound for the computational cost in terms of the desired accuracy, Tol_l has to be related to $N_l(\omega)$.

Assumption 3.5. For all $l = 1, \dots, L$ and each $\omega \in \Omega$, the dimension $N_l(\omega)$ of the ansatz space $S_l(\omega)$ providing the accuracy (9) satisfies

$$N_l(\omega) \le c_1 Tol_l^{-s} \tag{18}$$

with positive constants c_1, s independent of Tol_l, $N_l(\omega)$, and $\omega \in \Omega$.

Now we are ready to state an upper bound for the computational cost for the evaluation of $\mathbb{E}^L[u_L]$ in terms of the desired accuracy Tol. The proof is carried out along the lines of similar results in [18, 28].

THEOREM 3.2. *Let Assumptions* 2.1 *and* 3.1–3.5 *hold. Then there are numbers of samples M_l, $l = 1, \dots, L$, such that the inexact pathwise MLMC Galerkin approximation $\mathbb{E}^L[\tilde{u}_L]$ satisfies the error estimate*

$$\left\| \mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L] \right\|_{L^2(\Omega;H)} \le Tol \tag{19}$$

and can be evaluated with computational cost bounded by

$$C(L + |\log(Tol_1)|)^{\mu} Tol_1^{-s} L^{c_s} Tol^{-\max\{2,s\}} \quad \text{with} \quad \begin{cases} c_s = 2 & for \quad s = 2, \\ c_s = 0 & for \quad s \neq 2, \end{cases} \tag{20}$$

and a constant C only depending on c_0, c_1, q, s, μ, and $V[u]$.

Proof. Utilizing Assumptions 3.4 and 3.5, the computational cost for the evaluation of $\mathbb{E}^L[\tilde{u}_L]$ is bounded by

$$\begin{aligned}
&c_0 \sum_{l=1}^{L} \sum_{i=1}^{M_l} ((1 + \log(N_{l,i}(\omega)))^{\mu} N_{l,i}(\omega) + (1 + \log(N_{l-1,i}(\omega)))^{\mu} N_{l-1,i}(\omega_{li})) \\
&\quad \le c_0 c_1 \sum_{l=1}^{L} M_l \left((1 + \log(c_1 Tol_l^{-s}))^{\mu} Tol_l^{-s} + (1 + \log(c_1 Tol_{l-1}^{-s}))^{\mu} Tol_{l-1}^{-s} \right) \\
&\quad \le c_0 c_1 (1 + q^s) \sum_{l=1}^{L} M_l \left(1 + \log(c_1 q^{-s(l-1)} Tol_1^{-s}) \right)^{\mu} Tol_l^{-s} \\
&\quad \le c(L + |\log(Tol_1)|)^{\mu} \sum_{l=1}^{L} M_l Tol_l^{-s}
\end{aligned}$$

with a constant c depending on c_0, c_1, s, μ, and q. Hence, the desired upper bounds for the computational cost will follow from corresponding upper bounds for

$$\sum_{l=1}^{L} M_l Tol_l^{-s}.$$

We always select M_1 to be the smallest integer such that

$$M_1 \ge 12 \left(\tfrac{1}{4} Tol_1^2 + V[u] \right) Tol^{-2}, \tag{21}$$

so that the first term in the error estimate (16) is bounded by $\frac{1}{4} Tol^2$. The choice of the other M_l, $l = 2, \dots, L$, will depend on s.

Let us first consider the case $s < 2$. We choose the numbers of samples M_l to be the smallest integers such that

$$M_l \ge C_1 q^{\frac{s+2}{2}(l-1) + 2(1-L)}, \quad l = 2, \dots, L, \tag{22}$$

denoting $C_1 = 2(1+q^{-1})^2(1-q^{\frac{2-s}{2}})^{-1}$. Inserting any M_l, $l = 1,\dots,L$, with the properties (21) and (22) into the error estimate (16), we get

$$\left\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\right\|^2_{L^2(\Omega;H)} \le \tfrac{3}{4}Tol^2 + \tfrac{1}{2}Tol^2\left(1+q^{-1}\right)^2 C_1^{-1}\sum_{l=1}^{L-1} q^{\frac{2-s}{2}l} < Tol^2$$

by exploiting the convergence of geometric series. As we have chosen the smallest integers with the properties (21) and (22), we can exploit $q^{2(1-L)} = Tol_1^2 Tol^{-2}$, $Tol_l^{-s} = q^{-s(l-1)}Tol_1^{-s}$, $l = 2,\dots,L$, and similar arguments as above to obtain

$$\begin{aligned}\sum_{l=1}^{L} M_l Tol_l^{-s} &\le \left(12\left(\tfrac{1}{4}Tol_1^2 + V[u]\right)Tol_1^{-s} + C_1\sum_{l=2}^{L} q^{\frac{2-s}{2}(l-1)}\right)Tol^{-2} + \sum_{l=1}^{L} Tol_l^{-s}\\ &\le cTol_1^{-s}Tol^{-2}\end{aligned}$$

with a positive constant c depending on $s < 2$, q, and $V[u]$.

We now consider other values of s. The numbers of samples M_l are chosen to be the smallest integers such that

$$M_l \ge C_2 L q^{2(l-L)}, \quad l = 2,\dots,L, \tag{23}$$

with $C_2 = 2(1+q^{-1})^2$ for $s = 2$ and such that

$$M_l \ge C_3 q^{\frac{s+2}{2}(l-L)}, \quad l = 2,\dots,L, \tag{24}$$

with $C_3 = 2(1+q^{-1})^2(1-q^{\frac{s-2}{2}})^{-1}$ for $s > 2$. The same arguments as above then provide the desired bounds for accuracy and computational cost. ∎

Observe that the logarithmic term in Assumption 3.4 is reflected by the logarithmic term $(L + |\log(Tol_1)|)^\mu$ in the computational cost.

For $L = 1$, the approximation $\mathbb{E}[\tilde{u}_L]$ reduces to an inexact version of the classical MC method. Theorem 3.1 then implies that the error estimate (19) holds for

$$M \ge \tfrac{3}{2} + 6V[u]Tol^{-2}$$

with $M = M_1$ and $Tol = Tol_1$. The corresponding computational cost is bounded by

$$C(1 + |\log(Tol)|)^\mu Tol^{-(2+s)}$$

with C depending on c_0, c_1, s, μ, q, and $V[u]$, which indicates that, up to initial tolerance and logarithmic terms, the MLMC method is by a factor of $Tol^{-\min\{2,s\}}$ faster than the classical single level version.

4. Multilevel Monte Carlo finite element methods. We consider problem (1) with the symmetric bilinear form

$$a(\omega; v, w) = \int_D \alpha(x,\omega)\nabla v(x)\cdot\nabla w(x)\,dx \tag{25}$$

and the linear functional

$$\ell(\omega; v) = \int_D f(x,\omega)v(x)\,dx, \tag{26}$$

both defined on the Sobolev space $H = H_0^1(D)$ of weakly differentiable functions on a bounded Lipschitz domain $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, equipped with the norm

$$\|v\|_H = \left(\sum_{i=1}^{d} \left\| \tfrac{\partial}{\partial x_i} v \right\|_{L^2(D)}^2 \right)^{1/2}.$$

The closed convex set $K \in H$ of admissible solutions is given by

$$K = \{v \in H \mid v(x) \geq 0 \text{ a.e. in } D\}. \tag{27}$$

We impose the following assumptions on the random coefficient α and on the random right-hand side f.

Assumption 4.1. The random diffusion coefficient α and the right-hand side f are strongly measurable mappings $\Omega \ni \omega \mapsto \alpha(\cdot, \omega) \in L^\infty(D)$ and $\Omega \ni \omega \mapsto f(\cdot, \omega) \in L^2(D)$ with the properties

$$0 < \alpha_- \leq \alpha(x, \omega) \leq \alpha_+ < \infty \quad \text{a.e. in } D \times \Omega, \tag{28}$$

and $f \in L^2(\Omega; L^2(D))$.

These assumptions imply Assumption 2.1 and thus the existence and uniqueness of pathwise solutions $u(\omega)$ of (1) and $u \in L^2(\Omega; H)$. Note that uniform coercivity (28) can be replaced by weaker conditions (cf., e.g., [40]).

On the background of the general results from section 3 we now concentrate on MLMC finite element methods for the numerical approximation of the expectation $\mathbb{E}[u]$. Single level versions are obtained for the special case $L = 1$.

4.1. Uniform refinement. We assume for simplicity that D has a polygonal (polyhedral) boundary and consider the hierarchy of shape regular, conforming, quasiuniform partitions $\mathcal{T}^{(k)}$, $k \in \mathbb{N}$, of D into simplices as obtained by successive uniform refinement of a given, intentionally coarse, initial partition $\mathcal{T}^{(1)}$. (We will also assume that $\mathcal{T}^{(1)}$ is sufficiently fine in a sense to be specified later.)

Then

$$h_k = \max_{t \in \mathcal{T}^{(k)}} \operatorname{diam}(t) = 2^{-k} h_1, \qquad k \in \mathbb{N},$$

and the associated finite element spaces

$$S^{(k)} = \{v \in H \mid v|_t \text{ is affine } \forall t \in \mathcal{T}^{(k)}\}, \qquad k \in \mathbb{N}, \tag{29}$$

form a hierarchy of subspaces of H. We consider the pathwise approximations $u^{(k)}(\omega) \in K^{(k)} = S^{(k)} \cap K$ characterized by

$$a(\omega; u^{(k)}(\omega), v - u^{(k)}(\omega)) \geq \ell(\omega; v - u^{(k)}) \quad \forall v \in K^{(k)}, \qquad \omega \in \Omega. \tag{30}$$

Assumption 4.2. The spatial domain D is convex and the random coefficient α is a measurable map $\Omega \ni \omega \mapsto \alpha(\cdot, \omega) \in C^1(\bar{D})$ with the property $\alpha \in L^\infty(\Omega; C^1(\bar{D}))$.

The following discretization error estimate is a direct consequence of [40, Proposition 4.2].

THEOREM 4.1. *Let Assumptions* 4.1 *and* 4.2 *hold. Then the error estimate*

$$\|u - u^{(k)}\|_{L^2(\Omega;H)} \leq C_0 h_k \tag{31}$$

holds with a positive constant C_0 that is independent of h_k, $k \in \mathbb{N}$.

We make sure that $\mathcal{T}^{(1)}$ is fine enough to guarantee

$$\|u - u^{(1)}\|_{L^2(\Omega;H)} \leq \tfrac{1}{2\sqrt{2}} Tol_1 \tag{32}$$

by selecting h_1 such that $C_0 h_1 \leq \frac{1}{2\sqrt{2}} Tol_1$ and define a uniform MLMC hierarchy in the sense of (7) according to

$$S_l(\omega) = S^{(r(l-1)+1)}, \quad K_l = S_l(\omega) \cap K, \quad l = 1, \dots, L, \qquad \omega \in \Omega. \tag{33}$$

Then Assumption 3.1 is trivially satisfied and Theorem 4.1 implies the accuracy Assumption 3.2 by choosing $r \in \mathbb{N}$ such that $2^{-r} \leq q$. Furthermore, Assumption 3.3 can be satisfied by sufficiently many steps of any iterative solver for elliptic variational inequalities that converges uniformly in ω and consists of basic arithmetic or max operations, thus preserving measurability (cf., e.g., [20, 31, 39, 42, 46]). Then, by Theorem 3.1, the resulting uniform, inexact MLMC finite element approximation $\mathbb{E}^L[\tilde{u}_L]$ with sufficiently large numbers of MC samples M_l on each level satisfies the desired error estimate

$$\big\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\big\|_{L^2(\Omega;H)} \leq Tol. \tag{34}$$

It is well-known (cf. [40, section 4.5], [5, Corollary 4.1]) that standard monotone multigrid (STDMMG) methods [39, 42] satisfy Assumption 3.4 with $\mu = 4$ in $d = 1$ space dimension, with $\mu = 5$ in $d = 2$ space dimensions, and a suitable constant c_0. In spite of computational evidence, no theoretical justification of mesh-independent convergence rates seem to be available for $d = 3$. Finally, utilizing again Theorem 4.1, we find that Assumption 3.5 holds with $s = d$, because the dimension N_l of S_l is bounded by $h_{r(l-1)+1}^{-d}$ and thus by Tol_l^{-d} up to a constant c_1. Hence, Theorem 3.2 implies the following result on the efficiency of uniform MLMC finite element methods.

THEOREM 4.2. *Let Assumptions* 4.1 and 4.2 *and* (32) *hold, and let STDMMG be used for the iterative solution of the pathwise discretized obstacle problems of the form* (8).

Then there are M_l, $l = 1, \dots, L$, such that the resulting uniform MLMC finite element method provides an approximation $\mathbb{E}^L[\tilde{u}_L]$ with prescribed accuracy (34) *at computational cost bounded by*

$$C(L + d|\log Tol_1|)^{\mu} Tol_1^{-d} L^{c_s} Tol^{-\max\{2,d\}} \quad \textit{with} \quad \begin{cases} c_s = 0,\ \mu = 4 & \textit{for} \quad d = 1, \\ c_s = 2,\ \mu = 5 & \textit{for} \quad d = 2, \end{cases}$$

and a constant C depending only on c_0, c_1, q, and $V[u]$.

In fact, one could chose M_1 according to (21) and M_l, $l = 2, \dots, L$, according to (22) and (23) for $d = 1$ and $d = 2$, respectively.

The number of refinements in (33) can be defined a priori for all $\omega \in \Omega$. Hence, Theorem 4.2 is not new, but just a slightly enhanced version, e.g., of [40, Theorem 4.10]. Assuming that for all $k \in \mathbb{N}$ and each $\omega \in \Omega$ there is an a posteriori error estimate $\eta^{(k)}(\omega)$ satisfying

$$\|u(\omega) - u^{(k)}(\omega)\|_H \leq \eta^{(k)}(\omega),$$

a priori uniform refinement could be replaced by a posteriori uniform refinement with possibly different mesh sizes for different $\omega \in \Omega$. This approach can be regarded as a special case of a posteriori adaptive refinement presented in the next subsection.

4.2. Adaptive refinement. We consider a sequence of nested finite element spaces $S^{(k)}(\omega)$ associated with a corresponding sequence of partitions $\mathcal{T}^{(k)}(\omega)$, $k \in \mathcal{N}$, which, for each fixed $\omega \in \Omega$, is obtained by successive adaptive refinement of the given fixed initial triangulation $\mathcal{T}^{(1)}(\omega) = \mathcal{T}^{(1)}$. Let $\mathcal{T}^{(1)}$ be fine enough to provide the accuracy (32) and we set

$$S_1(\omega) = S^{(1)}, \qquad \omega \in \Omega. \tag{35}$$

For each fixed $\omega \in \Omega$ we apply a pathwise adaptive refinement providing a hierarchy of subspaces $S^{(k)}(\omega)$ and corresponding approximations $u^{(k)}(\omega)$. We assume convergence of the pathwise adaptive scheme controlled by an a posteriori error estimator.

Assumption 4.3. For all $k \in \mathbb{N}$ and for each fixed $\omega \in \Omega$ we have

$$\|u(\omega) - u^{(k)}(\omega)\|_H \le C_{est}\eta^{(k)}(\omega) \quad \text{and} \quad \eta^{(k)}(\omega) \xrightarrow{k\to\infty} 0 \tag{36}$$

with an a posteriori error estimator $\eta^{(k)}(\omega)$ and positive constant C_{est} independent of ω.

For each fixed $\omega \in \Omega$, there are existing adaptive algorithms based on local error indicators and corresponding a posteriori error estimates $\eta^{(k)}(\omega)$ that provide convergence (36); see, e.g., Siebert and Veeser [43], Braess, Carstensen, and Hoppe [15, section 5], or Carstensen [16]. The constant C_{est} in these algorithms usually depends on the initial triangulation $\mathcal{T}^{(1)}$ and on the ellipticity constants α_-, α_+.

We now define the hierarchy of subspaces for each $\omega \in \Omega$ according to

$$S_l(\omega) = S^{(k_l(\omega))}(\omega), \quad l = 2, \dots, L, \tag{37}$$

where $k_l(\omega)$ is the smallest natural number such that

$$\|u(\omega) - u^{(k_l(\omega))}(\omega)\|_H \le \tfrac{1}{2\sqrt{2}} Tol_l \tag{38}$$

and Tol_l is chosen according to (6). This definition makes sense, because $k_l(\omega) < \infty$ holds pointwise for each fixed $\omega \in \Omega$ by Assumption 4.3. Note that $k_l(\omega)$ might not be uniformly bounded in $\omega \in \Omega$. We assume that adaptive refinement and the accuracy criterion (38) preserve measurability.

Assumption 4.4. For all $l = 1, \dots, L$ the set-valued map $\Omega \ni \omega \mapsto S_l(\omega) \in H$ is measurable.

A rigorous investigation of sufficient conditions for measurability of $\omega \to S^{(k)}(\omega)$ and $\omega \to S_l(\omega)$ would exceed the scope of this presentation and is therefore postponed to a separate publication.

Assumption 4.4 clearly implies Assumption 3.1 while the initial condition (32) and the accuracy criterion (38) provide Assumption 3.2.

Assumption 3.3 can be satisfied by sufficiently many steps of any iterative solver for elliptic variational inequalities that converges uniformly in ω and consists of basic arithmetic or max operations, thus preserving measurability (cf., e.g., [20, 31, 39, 42, 46]).

Like in the uniform case, Assumption 3.4 can be satisfied by STDMMG methods [39, 42] with $\mu = 4$ in $d = 1$ space dimension and $\mu = 5$ in $d = 2$ space dimensions with a suitable constant c_0.

Now, instead of the regularity Assumption 4.2, we require that pathwise adaptive refinement provides quasioptimal meshes uniformly in $\omega \in \Omega$.

Assumption 4.5. For all $l = 1, \dots, L$ and each $\omega \in \Omega$, the dimension $N_l(\omega)$ of the finite element space $S_l(\omega)$ defined in (35) and (37) satisfies

$$N_l(\omega) \le c_1 Tol_l^{-d} \tag{39}$$

with a positive constant c_1 independent of Tol_l, $N_l(\omega)$, and $\omega \in \Omega$.

For fixed $\omega \in \Omega$ and $K = H$, the quasioptimality condition (39) has been established for a variety of adaptive refinement strategies with a constant $c_1(\omega)$ (cf., e.g., [11, 44, 41]). Uniform upper bounds for $c_1(\omega)$ as required in Assumption 4.5 are observed in the numerical experiments to be presented in the next section. Theoretical validation will be the subject of future research.

Now the following convergence result is a direct consequence of Theorem 3.2.

THEOREM 4.3. *Let Assumptions* 4.1 and 4.3–4.5 *and* (32) *hold. Then there are* M_l, $l = 1, \dots, L$, *such that the adaptive MLMC finite element method based on the multilevel hierarchy defined in* (37) *provides an approximation* $\mathbb{E}[\tilde{u}_L]$ *with prescribed accuracy* (34) *at computational cost bounded by*

$$C(L + d|\log Tol_1|)^{\mu} Tol_1^{-d} L^{c_s} Tol^{-\max\{2,d\}} \quad \textit{with} \quad \begin{cases} c_s = 0,\ \mu = 4 & \textit{for} \quad d = 1, \\ c_s = 2,\ \mu = 5 & \textit{for} \quad d = 2, \end{cases}$$

and a constant C *depending only on* c_0, c_1, q, *and* $V[u]$.

In fact, one could chose M_1 according to (21) and M_l, $l = 2, \dots, L$, according to (22) and (23) for $d = 1$ and $d = 2$, respectively.

5. Numerical experiments. In this section we investigate the adaptive MLMC finite element approach presented in the preceding sections from a numerical perspective. We use the algorithm proposed by Giles [29, Algorithm 1] (see also [28]). Here, the increment of the number of levels is associated with uniform mesh refinement for uniform MLMC and an update of the stopping criterion for adaptive MLMC to be specified later. We slightly modified the computation of the optimal number of realizations on each level by replacing the cost of an individual realization by the average of the cost of all realizations on the same level. In our computations, we used a minimal number $M_{\min}$ of samples setting $M_{\min} = 100$ for the Poisson problem (cf. subsection 5.1) and $M_{\min} = 50$ for the obstacle problem (cf. subsection 5.2).

The initial accuracy condition (32) is addressed by formally setting

$$Tol_1 = 2\sqrt{2} C_{est} \|\eta^{(1)}\|_{L^2(\Omega)} \tag{40}$$

with the $L^2(\Omega)$-norm approximated by an MC method with 1000 samples. We choose Tol_l according to (6) with $q = \frac{1}{2}$. The accuracy criterion (38) is replaced by the approximation

$$\eta^{(k_l(\omega))}(\omega) \le \tfrac{1}{2\sqrt{2} C_{est}} Tol_l = q^{l-1} \|\eta^{(1)}\|_{L^2(\Omega)}, \tag{41}$$

which is used as stopping criterion on each level in adaptive MLMC. Note that the unknown constant C_{est} does not appear in our computations. Both uniform and adaptive MLMC terminate once the stopping criterion in Giles' algorithm is met.

Pathwise adaptive refinement is performed as suggested by Siebert and Veeser [43] with error indicators $\eta_t(\omega)$ given by local contributions to the hierarchical error estimator according to [49, Theorem 3.5]. Here, the exact finite element solution is replaced by an approximation provided by an iterative method to be described below.

In the unconstrained case $K = H$, this approach is reducing to the classical hierarchical error estimation (cf., e.g., [14, 22] or [23, section 6.1.4]). Note that the error is estimated in the energy norm. We use Dörfler marking [24] with $\theta = 0.4$ for the Poisson problem (cf. subsection 5.1) and $\theta = 0.2$ for the obstacle problem (cf. subsection 5.2) together with local "red" mesh refinement [6, 9, 13] with hanging nodes [30, section 3.1]. Implementation is carried out in the finite element software environment DUNE [12] involving the `dune-subgrid` module [32] for the evaluation of the sum of different approximate evaluations of $u_{l,i}(\omega)$ on different grids.

Discretized variational inequalities of the form (30) are solved iteratively by truncated nonsmooth Newton multigrid methods (TNNMG) [30, 31] with nested iterations, because TNNMG is easier to implement and usually converges faster than STDMMG [30]. Numerical experiments (see, e.g., [40, section 5]) also indicate that TNNMG satisfies Assumption 3.4 with $\mu = 0$. Note that both STDMMG and TNNMG reduce to classical multigrid with Gauss–Seidel smoothing in the unconstrained case $K = H$. The accuracy condition (10) is replaced by the uniform stopping criterion

$$\|u_{\nu+1}^{(k)} - u_{\nu}^{(k)}\|_H \leq \tfrac{1}{2\sqrt{2}}\sigma_{alg} Tol_l$$

with $u_{\nu}^{(k)}$ denoting the νth iterate and a safety factor $\sigma_{alg} = 0.001$ accounting for estimating the algebraic error $\|u^{(k)} - u_{\nu}^{(k)}\|_H$ by $\|u_{\nu+1}^{(k)} - u_{\nu}^{(k)}\|_H$. In view of the abovementioned optimal convergence properties of TNNMG, the cost for the evaluation of $\tilde{u}_l(\omega) \in S_l(\omega)$ is set to the corresponding number of unknowns $N_l(\omega) = \dim S_l(\omega)$, i.e.,

$$cost(\tilde{u}_l(\omega)) = N_l(\omega).$$

In light of (17), the computational cost for the adaptive MLMC method with L levels is then given by

$$cost_L = \sum_{l=1}^{L} \sum_{i=1}^{M_l} N_{l,i}(\omega), \tag{42}$$

which reduces to $cost_L = \sum_{l=1}^{L} N_l M_l$ in case of uniform refinement.

5.1. Poisson equation with random right-hand side. We consider the Poisson problem

$$u(\omega) \in \left\{ w \in H^1(D) \mid w|_{\partial D} = g(\omega) \right\} : \quad a(\omega; u(\omega), v) = \ell(\omega; w) \qquad \forall v \in H_0^1(D) \tag{43}$$

with $D = (-1, 1)^2$ in $d = 2$ space dimensions, the bilinear form

$$a(\omega; v, w) = \int_D \nabla v(x) \cdot \nabla w(x)\, dx, \quad v, w \in H, \tag{44}$$

the right-hand side

$$\ell(\omega; v) = \int_D f(x, \omega) v(x)\, dx, \quad v \in H, \tag{45}$$

with uncertain source term

$$f(x, \omega) = e^{-\beta|x - Y(\omega)|^2} \left(4\beta^2 |x - Y(\omega)|^2 - 4\beta\right), \tag{46}$$

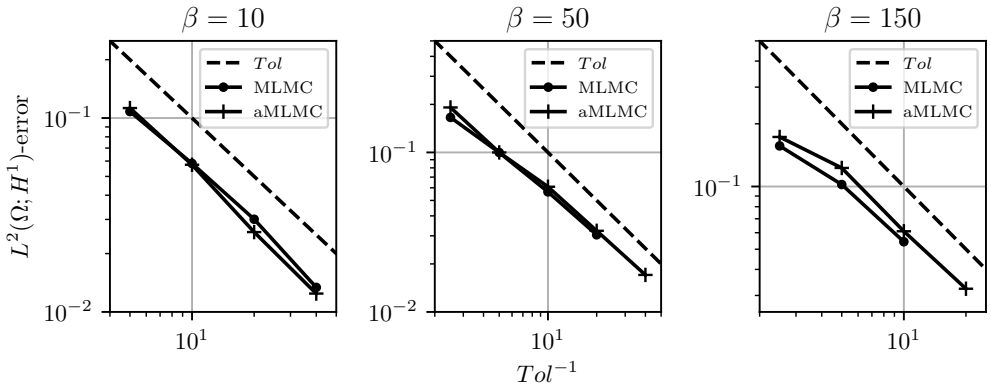

FIG. 1. *Error achieved by uniform and adaptive MLMC over the inverse of required accuracy Tol for the Poisson problem.*

and uncertain, inhomogeneous boundary conditions

$$g(x,\omega) = e^{-\beta|x-Y(\omega)|^2}, \quad x \in \partial D.$$

Here, β is a positive constant and $Y(\omega) = (Y_1(\omega), Y_2(\omega))^T$ is a random vector whose components are uniformly distributed random variables $Y_1, Y_2 \sim \mathcal{U}(-0.25, 0.25)$. For each $\omega \in \Omega$ a pathwise solution of (43) is given by

$$u(x,\omega) = e^{-\beta|x-Y(\omega)|^2}, \quad x \in D. \tag{47}$$

As Assumption 4.1 is satisfied, this solution is unique and we have spatial regularity in the sense that $u \in L^2(\Omega; H^2(D))$ (cf. Assumption 4.2). However, $u(\omega)$ exhibits a peak at $(Y_1(\omega), Y_2(\omega)) \in D$ that becomes more pronounced with increasing β, thus leading to larger constants C_0 in the uniform error estimate (31).

We will compare the performance of MLMC finite element methods based on uniform and adaptive refinement, as presented in the preceding section 4, for $\beta = 10$, 50, 150. The initial partition $\mathcal{T}^{(1)}$ is obtained by applying four uniform refinement steps to the partition of the unit square $\overline{D}$ into two congruent triangles with right angles at $(1,-1)$ and $(-1,1)$.

Figure 1 illustrates the convergence properties of uniform and adaptive MLMC methods for the different values of β by showing the actually achieved error over the inverse of the required tolerance Tol. Here, the error $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H^1(D))}$ is approximated by an MC method utilizing $M = 5$ independent realizations $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$. For all values of β, both uniform and adaptive MLMC match the required accuracy Tol as indicated by the dotted line, thus nicely confirming our theoretical results (cf. Theorems 4.1 and 4.3) also in this slightly more general case of random boundary conditions. Due to limited memory resources the accessible accuracy of uniform MLMC is exceeded by adaptive MLMC for $\beta = 50, 150$

We now investigate the corresponding computational effort in terms of required number of samples and mesh size. Figure 2 shows the average numbers of optimal MLMC samples M_l (sometimes smaller than $M_{\min}$) over the corresponding levels $l = 1, \ldots, L$ for different values of β and Tol. The average is taken over the $M = 5$ realizations of $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$. It is interesting that the number of samples required for adaptive MLMC is always smaller than for uniform MLMC and that the

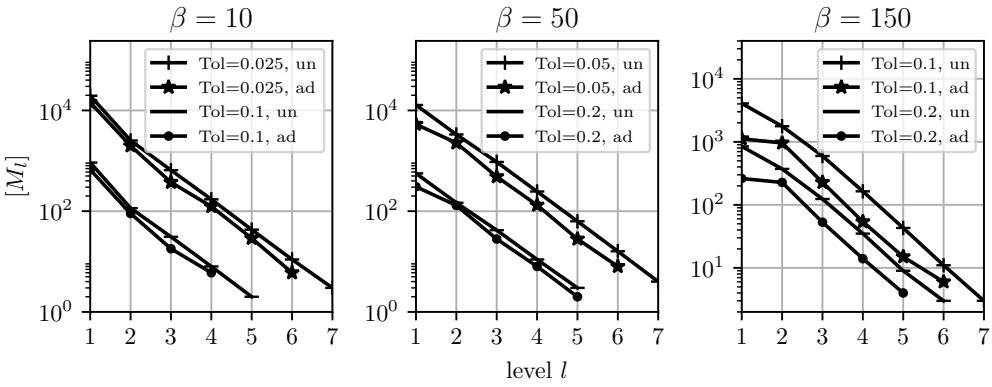

FIG. 2. *Average optimal number of samples over levels for uniform (un) and adaptive (ad) MLMC, different values of Tol, and the Poisson problem.*

TABLE 1
Average number of unknowns on different levels for the Poisson problem.

l	1	2	3	4	5	6
Uniform	289	1089	4225	16641	66049	263169
Adaptive, $\beta = 10$	289	965	3339	11719	56087	218507
Adaptive, $\beta = 50$	289	508	1017	5701	16901	49895
Adaptive, $\beta = 150$	289	385	929	2730	6938	19606

difference becomes larger for larger β. Moreover, adaptive MLMC often requires less levels than uniform MLMC.

Table 1 reports on the average mesh sizes or, equivalently, the average of the number of the unknowns $N_{l,i}(\omega)$, $i = 1, \dots, M_l$, on the levels $l = 1, \dots, 7$ for uniform and adaptive MLMC up to tolerances 0.025, 0.05, and 0.1 for $\beta = 10$, 50, and 100, respectively. Note that adaptive MLMC reached the desired tolerances already on level $L = 6$. While for $\beta = 10$ the corresponding uniform and adaptive mesh sizes stay relatively close to each other, the mesh sizes for adaptive MLMC for $\beta = 50$, 150 are considerably smaller than for uniform MLMC. Even though most of the work in MLMC methods is performed on coarser levels, this already indicates a gain of efficiency by adaptive mesh refinement.

Upper bounds of the computational cost of MLMC in terms of the desired accuracy Tol as stated in Theorem 3.2 strongly rely on Assumption 3.5 postulating $N_l(\omega) = \mathcal{O}(Tol_l^{-s})$. While, under suitable regularity conditions, Assumption 3.5 holds with $s = d$ for uniform MLMC, there is no theoretical evidence yet for adaptive MLMC. In order to check Assumption 3.5 for adaptive MLMC numerically, we adaptively computed approximations to realizations of $u_{l,i}(\omega)$, $i = 1, \dots, I = 1000$, up to the tolerance $\frac{1}{2\sqrt{2}C_{est}} Tol_l$ according to the stopping criterion (41) for $l = 1, \dots, 7$, and $\beta = 10$, 50, 150. Figure 3 displays the maximal required number of unknowns $N_{l,\max} = \max_{i=1,\dots,I} N_{l,i}(\omega)$ over the the number of levels $l = 1, \dots, 7$. We observe that $\log(N_{l,\max})$ grows like $2\log(q)(l-1)$ (dotted line) or, equivalently, $N_{l,\max} = \mathcal{O}(Tol_l^{-2})$ for all three values of β. This indicates that adaptive MLMC satisfies Assumption 3.5 with $s = d = 2$.

On this background, we expect from Theorem 3.2 that the computational cost both of uniform and adaptive MLMC should asymptotically behave like $\mathcal{O}(Tol^{-2})$. Figure 4 shows the average of $cost_L$, as defined in (42), over the inverse of the required

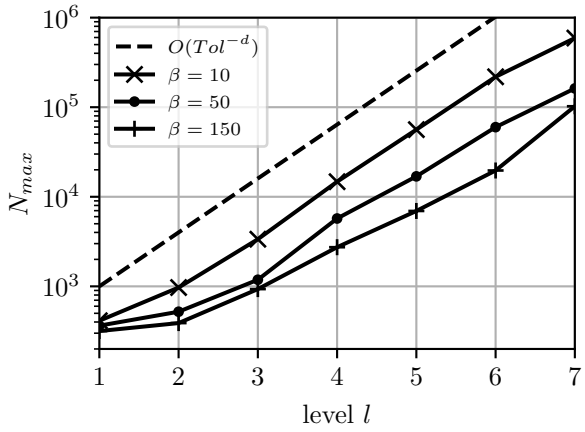

FIG. 3. *Number of unknowns providing the accuracy $\mathcal{O}(Tol_l)$ over levels l for the Poisson problem.*

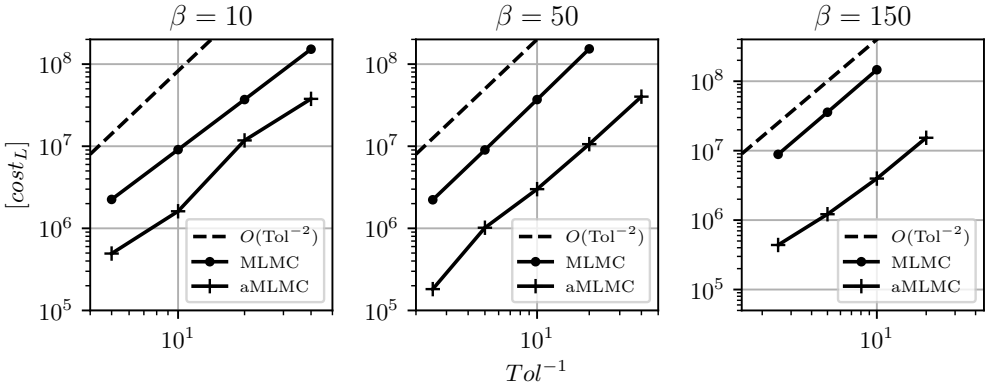

FIG. 4. *Average computational cost of uniform and adaptive MLMC over the inverse of required accuracy Tol for the Poisson problem.*

accuracy Tol together with the expected asymptotic behavior (dotted line). As in Figure 1, the average is taken over the $M = 5$ realizations of $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$. Observe that adaptive MLMC always outperforms uniform MLMC and the gain is increasing with increasing β. Though the simple model of computational cost (42) is frequently used, it obviously ignores a posteriori error estimation, mesh handling, interpolation, etc., which does occur in adaptive MLMC but not in the uniform case. We therefore complement our considerations by a comparison of the overall run time on the machine with 3.3 a GHz Intel Xeon E3-1245 processor with 7.8 GBytes of RAM for different tolerances Tol and different values of β. We found that the overall run time to reach the tolerance, $Tol = 0.025$, 0.05, and 0.1, by uniform MLMC was improved by a factor of 1.1, 3.2, and 4.6 by adaptive MLMC for $\beta = 10$, 50, and 150, respectively. These experiments confirm that uniform MLMC is preferable for sufficiently smooth problems while, even without specific software optimization, adaptive MLMC can substantially reduce the computational cost in the presence of random singularities.

5.2. Obstacle problem with random diffusion coefficient and right-hand side. We consider an elliptic variational inequality of the form (1) with $D = (0, 1)$ in

$d = 1$ space dimension,

$$K = \{v \in H \mid v(x) \geq 0 \text{ a.e. in } D\} \subset H, \quad H = H_0^1(D),$$

the bilinear form

$$a(\omega; v, w) = \int_D \alpha(x, \omega) \nabla v(x) \cdot \nabla w(x)\, dx, \quad v, w \in H, \tag{48}$$

with random diffusion coefficient

$$\alpha(x, \omega) = 1 + \frac{\cos x^2}{10} Y_1(\omega) + \frac{\sin x^2}{10} Y_2(\omega), \tag{49}$$

and the right-hand side

$$\ell(\omega; v) = \int_D f(x, \omega) v(x)\, dx, \quad v \in H, \tag{50}$$

with random source term

$$f(x,\omega) = \begin{cases} -8e^{2(Y_1(\omega)+Y_2(\omega))} \left(a(x,\omega) \cdot (3x^2 - r^2)\right. & \\ \quad \left. + (x^2 - r^2)x^2 \left(-\frac{\sin x^2}{10} Y_1(\omega) + \frac{\cos x^2}{10} Y_2(\omega)\right)\right), & x > r, \\ 4r^2 e^{2(Y_1(\omega)+Y_2(\omega))} \left(a(x,\omega) \cdot (-1 - r^2 + x^2)\right. & \\ \quad \left. + (-2 - 2r^2 + x^2)x^2 \left(-\frac{\sin x^2}{10} Y_1(\omega) + \frac{\cos x^2}{10} Y_2(\omega)\right)\right), & x \leq r, \end{cases}$$

denoting

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

Here, $Y_1, Y_2 \sim \mathcal{U}(-1, 1)$ stand for uniformly distributed random variables. For each $\omega \in \Omega$ a solution of the corresponding pathwise problem (1) is given by

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

As Assumption 4.1 is satisfied, this solution is unique and we have $u \in L^2(\Omega; H)$.

We will compare the numerical behavior of MLMC finite element methods with uniform and adaptive spatial mesh refinement as presented in section 4. The initial partition $\mathcal{T}^{(1)}$ of $\overline{D} = [0, 1]$ consists of sixteen closed intervals with length $1/16$.

Figure 5 shows the error $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H^1(D))}$ of uniform and adaptive MLMC over Tol^{-1}. As in the previous numerical experiment, the exact error $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega;H^1(D))}$ is approximated by an MC method utilizing $M = 5$ independent realizations $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$. As expected from Theorems 4.1 and 4.3, both for uniform and adaptive MLMC the error is bounded by the prescribed tolerance Tol indicated by the dotted line. Adaptive MLMC appears to be slightly more accurate than the uniform version.

Next, we consider the required number of samples and mesh size. The average optimal number of MLMC samples M_l over the corresponding levels $l = 1, \ldots, L$ are shown in Figure 6 for different values of Tol. Again, the numbers of samples for adaptive MLMC are slightly smaller than for the uniform method.

The average mesh size or, equivalently, the average of the number of unknowns $N_{l,i}(\omega)$, $i = 1, \ldots, M_l$ on the levels $l = 1, \ldots, 10$ for prescribed tolerance 0.00125 is

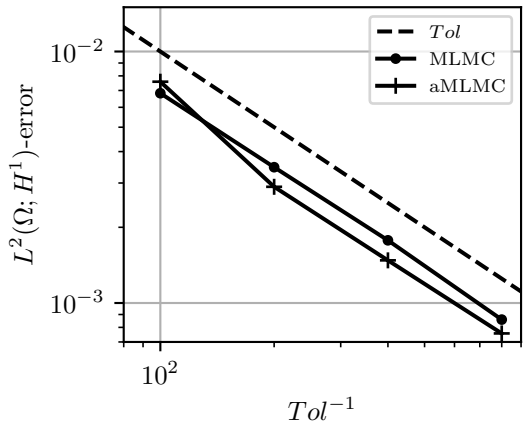

FIG. 5. *Error achieved by uniform and adaptive MLMC over the inverse of required accuracy Tol for the obstacle problem.*

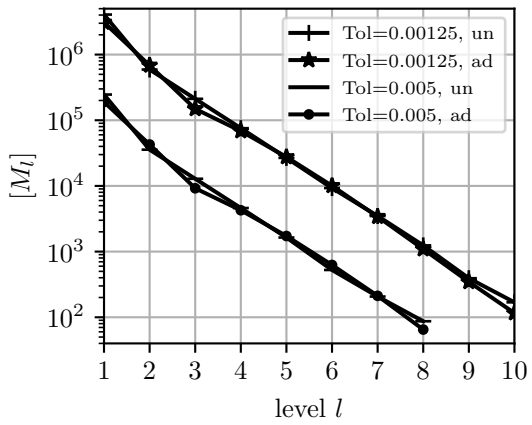

FIG. 6. *Average optimal number of samples over levels for uniform (un) and adaptive (ad) MLMC, different values of Tol, and the obstacle problem.*

TABLE 2
Average number of unknowns on different levels for the obstacle problem.

l	1	2	3	4	5	6	7	8	9	10
Uniform	17	33	65	129	257	513	1025	2049	4097	8193
Adapted	17	19	24	34	57	106	207	443	927	2150

reported in Table 2. The uniform mesh size on the final level $L = 10$ is about 3.8 times larger than for adaptive MLMC indicating the potential of the adaptive approach.

As the given data clearly satisfy Assumption 4.2, the general Assumption 3.5 holds true for uniform MLMC. Hence, Theorem 4.2 provides the upper bound $\mathcal{O}(Tol^{-2})$ for the computational cost of uniform MLMC. As corresponding theoretical evidence is still missing for adaptive MLMC, we check Assumption 3.5 numerically. To this end, we adaptively computed approximations to realizations of $u_{l,i}(\omega)$, $i = 1, \dots, I = 1000$, up to the tolerance $\frac{1}{2\sqrt{2}C_{est}} Tol_l$ according to the stopping criterion (41) for $l = 1, \dots, 12$. Figure 7 displays $N_{l,\max} = \max_{i=1,\dots,I} N_{l,i}(\omega)$ over the number of levels

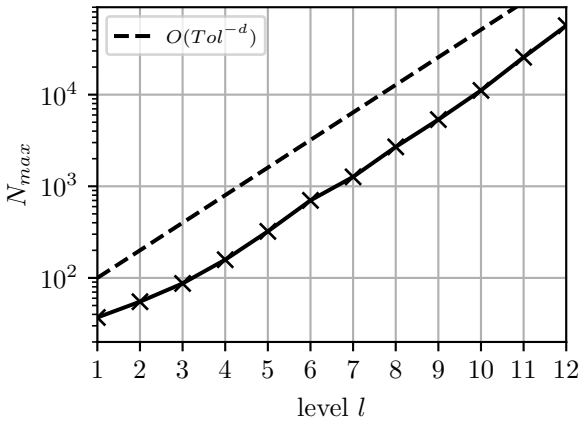

FIG. 7. *Number of unknowns providing the accuracy $\mathcal{O}(Tol_l)$ over levels l for the obstacle problem.*

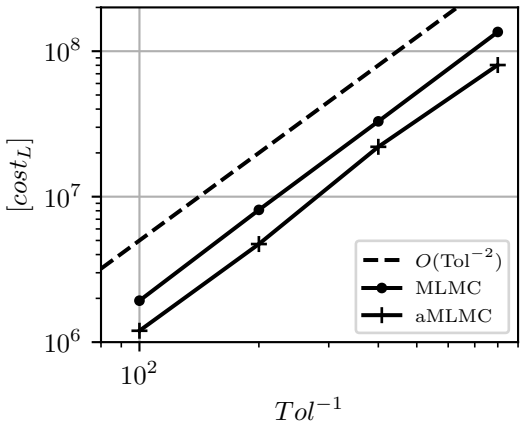

FIG. 8. *Average computational cost of uniform and adaptive MLMC over the inverse of required accuracy Tol for the obstacle problem.*

l. We observe that $\log(N_{l,\max})$ grows like $\log(q)(l-1)$ (dotted line) or, equivalently, $N_{l,\max} = \mathcal{O}(Tol_l^{-1})$ indicating that adaptive MLMC satisfies Assumption 3.5 with $s = d = 1$.

From Theorems 4.2 and 3.2, combined with numerical evidence of Assumption 3.5, we expect that the computational cost both of uniform and adaptive MLMC asymptotically behaves like $\mathcal{O}(Tol^{-2})$. This is confirmed by Figure 8, showing the average computational cost over the inverse of the required accuracy Tol together with the expected asymptotic behavior (dotted line). Again, the average is taken over the $M = 5$ realizations of $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$. We observe a gain of efficiency of adaptive MLMC by a factor of 1.75 as compared to the uniform version.

We also measured the overall run time on the machine with a 3.3 GHz Intel Xeon E3-1245 processor with 7.8 GBytes of RAM for the final tolerance $Tol = 0.00125$ and found that (for the given implementation) the overall run time is not improved by adaptive refinement.

Acknowledgments. The authors want to thank Robert Scheichl for stimulating discussions and the unknown referees for their thorough review and their valuable suggestions that significantly improved the paper.

REFERENCES

[1] A. Abdulle, A. Barth, and C. Schwab, *Multilevel Monte Carlo methods for stochastic elliptic multiscale PDEs*, Multiscale Model. Simul., 11 (2013), pp. 1033–1070.

[2] M. Ainsworth and J.T. Oden *A Posteriori Error Estimation in Finite Element Analysis*, Wiley, New York, 2011.

[3] I. Babuška, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Rev., 52 (2010), pp. 317–355.

[4] I. Babuška, R. Tempone, and G.E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825.

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

[6] R.E. Bank, A.H. Sherman, and A. Weiser, *Refinement algorithms and data structures for regular local mesh refinement*, in Scientific Computing: Application of Mathematics, and Computing to the Physical Sciences, R. Stepleman, ed.,, North Holland, Amsterdam, 1983, pp. 3–17.

[7] A. Barth, A. Lang, and C. Schwab, *Multilevel Monte Carlo method for parabolic stochastic partial differential equations*, BIT, 53 (2013), pp. 3–27.

[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, *Simplicial grid refinement: On Freudenthal's algorithm and the optimal number of congruence classes*, Numer. Math., 85 (2000), pp. 1–29.

[10] C. Bierig and A. Chernov, *Convergence analysis of multilevel Monte Carlo variance estimators and application for random obstacle problems*, Numer. Math., 130 (2015), pp. 579–613.

[11] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., 97 (2004), pp. 219–268.

[12] M. Blatt, A. Burchardt, A. Dedner, Ch. Engwer, J. Fahlke, B. Flemisch, Ch. Gersbacher, C. Graser, F. Gruber, Ch. Gruninger, D. Kempf, R. Klofkorn, T. Malkmus, S. M̈uthing, M. Nolte, M. Piatkowski, and O. Sander, *The distributed and unified numerics environment, version* 2.4, Archive of Numerical Software, 4 (2016), pp. 13–29.

[13] F. Bornemann, B. Erdmann, and R. Kornhuber, *Adaptive multilevel methods in three space dimensions*, Internat. J. Numer. Methods Engrg., 36 (1993), pp. 3187–3203.

[14] F. 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.

[15] D. Braess, C. Carstensen, and R.H.W. Hoppe, *Convergence analysis of a conforming adaptive finite element method for an obstacle problem*, Numer. Math., 107 (2007), pp. 455–471.

[16] C. Carstensen, M. Feischl, M. Page, and D. Praetorius, *Axioms of adaptivity*, Comput. Methods Appl. Math., 67 (2014), pp. 1195–1253.

[17] J. Charrier, R. Scheichl, and A. 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.

[18] K.A. Cliffe, M.B. Giles, R. Scheichl, and A.L. Teckentrup, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci., 14 (2011), pp. 3–15.

[19] N. Collier, A.-L. Haji-Ali, F. Nobile, E. von Schwerin, and R. Tempone, *A continuation multilevel Monte Carlo algorithm*, BIT, 55 (2015), pp. 399–432.

[20] R.W. Cottle and J.-S. Pang, *The Linear Complementarity Problem*, SIAM, Philadelphia, 1992.

[21] G. Detommaso, T. Dodwell, and R. Scheichl, *Continuous Level Monte Carlo and Sample-adaptive Model Hierarchies*, preprint, arXiv:1802.07539, 2018.

[22] P. Deuflhard, P. Leinen, and H. Yserentant, *Concepts of an adaptive hierarchical finite element code*, IMPACT Comput. Sci. Engrg., 1 (1989), pp. 3–35.

[23] P. Deuflhard and M. Weiser, *Adaptive Numerical Solution of PDEs*, Vol. 55, de Gruyter, Berlin, 2012.

[24] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.
[25] M. Eigel, Ch. Merdon, and J. Neumann, *An Adaptive Multilevel Monte Carlo Method with Stochastic Bounds for Quantities of Interest with Uncertain Data*, WIAS preprint 2060, 2015.
[26] D. Elfverson, F. Hellman, and A. Målqvist, *A multilevel Monte Carlo method for computing failure probabilities*, SIAM/ASA J. Uncertain. Quantif., 4 (2016), pp. 312–330.
[27] T. Gerstner and S. Heinz, *Dimension- and time-adaptive multilevel Monte Carlo methods*, in Sparse Grids and Applications, J. Garcke and M. Griebel, eds., Lect. Notes Comput. Sci. Eng. 88, Springer, New York, 2013, pp. 107–120.
[28] M.B. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.
[29] M.B. Giles, *Multilevel Monte Carlo methods*, Acta Numer., 24 (2015), pp. 259–328.
[30] C. Gräser, *Convex Minimization and Phase Field Models*, Ph.D. thesis, FU Berlin, 2011.
[31] C. Gräser and R. Kornhuber, *Multigrid methods for obstacle problems*, J. Comput. Math., 27 (2009), pp. 1–44.
[32] C. Gräser and O. Sander, *The dune-subgrid module and some applications*, Computing, 8 (2009), pp. 269–290.
[33] 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.
[34] J. Gwinner and F. Raciti, *On a class of random variational inequalities on random sets*, Numer. Func. Anal. Optim., 27 (2006), pp. 619–636.
[35] A.-L. Haji-Ali, F. Nobile, E.v. Schwerin, and R. Tempone, *Optimization of mesh hierarchies in multilevel Monte Carlo samplers*, Stoch. Partial Differ. Equ. Anal. Comput., 4 (2016), pp. 76–112.
[36] H. Hoel, E. Von Schwerin, A. Szepessy, and R. Tempone, *Implementation and analysis of an adaptive multilevel Monte Carlo algorithm*, Monte Carlo Methods Appl., 20 (2014), pp. 1–41.
[37] H. Hoel, E. Von Schwerin, A. Szepessy, and R. Tempone, *Adaptive multilevel Monte Carlo simulation*, in Numerical Analysis of Multiscale Computations, Springer, New York, 2012, pp. 217–234.
[38] D. Kinderlehrer and G. Stampaccia, *Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, 1980.
[39] R. Kornhuber, *Monotone multigrid methods for elliptic variational inequalities* I. Numer. Math., 69 (1994), pp. 167–184.
[40] R. Kornhuber, C. Schwab, and M.-W. Wolf, *Multilevel Monte Carlo finite element methods for stochastic elliptic variational inequalities*, SIAM J. Numer. Anal., 52 (2014), pp. 1243–1268.
[41] C. Kreuzer and K.G. Siebert, *Decay rates of adaptive finite elements with Dörfler marking*, Numer. Math., 117 (2011), pp. 679–716.
[42] J. Mandel, *A multilevel iterative method for symmetric, positive definite linear complementarity problems*, Appl. Math. Optim., 11 (1984), pp. 77–95.
[43] K. Siebert and A. Veeser, *A unilaterally constrained quadratic minimization with adaptive finite elements*, SIAM J. Optim., 18 (2007), pp. 260–289.
[44] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
[45] A. Teckentrup, R. Scheichl, M. Giles, and E. Ullmann, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numer. Math., 125 (2013), pp. 569–600.
[46] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, SIAM, Philadelphia, 1992.
[47] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley, New York, 1996.
[48] D. Xiu and G.E. Karniadakis, *The Wiener–Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2002), pp. 619–644.
[49] Q. Zou, A. Veeser, R. Kornhuber, and C. Gräser, *Hierarchical error estimates for the energy functional in obstacle problems*, Numer. Math., 117 (2011), pp. 653–677.