



MOX-Report No. 32/2026

**A Multilevel Monte Carlo Virtual Element Method for Uncertainty  
Quantification of Elliptic Partial Differential Equations**

Antonietti, P.F.; Bonizzoni, F.; Perugia, I.; Verani, M.

MOX, Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

[mox-dmat@polimi.it](mailto:mox-dmat@polimi.it)

<https://mox.polimi.it>

# A Multilevel Monte Carlo Virtual Element Method for Uncertainty Quantification of Elliptic Partial Differential Equations\*

Paola F. Antonietti<sup>a</sup>, Francesca Bonizzoni<sup>a</sup>, Ilaria Perugia<sup>b</sup>, Marco Verani<sup>a</sup>

<sup>a</sup> MOX, Dipartimento di Matematica, Politecnico di Milano, Milano, Italy

<sup>b</sup> Faculty of Mathematics, University of Vienna, Vienna, Austria

April 9, 2026

## Abstract

We introduce a Monte Carlo Virtual Element estimator based on Virtual Element discretizations for stochastic elliptic partial differential equations with random diffusion coefficients. We prove estimates for the statistical approximation error for both the solution and suitable linear quantities of interest. A Multilevel Monte Carlo Virtual Element method is also developed and analyzed to mitigate the computational cost of the plain Monte Carlo strategy. The proposed approach exploits the flexibility of the Virtual Element method on general polytopal meshes and employs sequences of coarser spaces constructed via mesh agglomeration, providing a practical realization of the multilevel hierarchy even in complex geometries. This strategy substantially reduces the number of samples required on the finest level to achieve a prescribed accuracy. We prove convergence of the multilevel method and analyze its computational complexity, showing that it yields significant cost reductions compared to standard Monte Carlo methods for a prescribed accuracy. Extensive numerical experiments support the theoretical results and demonstrate the efficiency of the proposed method.

**Keywords:** Multilevel Monte Carlo, Virtual Element Methods, Stochastic Partial Differential Equations, Uncertainty Quantification, Polytopal Meshes, Mesh Agglomeration

**AMS subject classification:** 65N30, 65N15, 60H15, 60H35, 65C30, 65C05

## 1 Introduction

Differential models with uncertain data are common in engineering, the social sciences, and the life sciences. Indeed, models' input data, such as coefficients or the source term, are often uncertain due to, e.g., intrinsic variability or measurement errors. A typical approach to handling such uncertainties is to model the uncertain data as random fields and analyze statistical quantities associated with solutions to the resulting stochastic PDEs. These quantities often include the expected value of the solution or functionals (quantity of interest, QoI) that depend on it. The Monte Carlo (MC) method is widely used to approximate solutions to stochastic PDEs by solving the corresponding deterministic problem repeatedly with independent realizations of the random input and computing statistical measures. However, this method can be impractical when each deterministic problem is computationally expensive, since the error decreases at a rate of  $O(M^{-1/2})$ , where  $M$  is the number of samples. To overcome this limitation, Multilevel Monte Carlo (MLMC) methods have been developed. The key idea is to combine many inexpensive coarse simulations with fewer costly fine ones, achieving an improved complexity–error balance. Our approach builds on the MLMC framework for infinite-dimensional integration introduced by Giles in the context of stochastic differential equations [16, 17]. It was then extended to elliptic PDEs with random coefficients in [5, 14, 13, 21] and more recently in [6, 20]. For a review of MLMC and its variants, see [18]. The effectiveness of MLMC depends on the availability of a sequence of coarser discretizations to approximate the corresponding deterministic problem. For example, in the Finite Element method, a sequence of coarser meshes is required for practical MLMC implementation. In complex geometries, constructing such a sequence of coarser (simplicial or hexahedral) meshes can be challenging. On the other

---

\*This research has been partially funded by the European Union (ERC, NEMESIS, project number 101115663). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This research was also funded in part by the Austrian Science Fund (FWF) 10.55776/F65. The present research is part of the activities of “Dipartimento di Eccellenza 2023-2027”. PA and MV also acknowledge MUR-PRIN/PNRR 2022 grant n. P2022BH5CB, funded by MUR. PFA, FB and MV are members of INdAM-GNCS.

hand, mesh agglomeration, which merges groups of fine-mesh elements into larger cells of arbitrary polygonal or polyhedral shape, provides a robust, automatic way to generate necessary coarser problems [4, 2, 15, 3]. This has driven the development of discretization techniques capable of handling general polytopal meshes. Among polytopal methods that have been developed in the last decade, in this paper we focus on the Virtual Element (VE) method that has been proposed in the seminal papers [8, 9, 10] and further successfully developed for a wide set of problems in computational sciences; see, e.g., the monographs [11] and [1], and the references therein.

In this work, we consider a stochastic elliptic PDE with random diffusion coefficients and aim to approximate both the expected value of the solution and that of a linear quantity of interest. We introduce and analyse a novel multilevel Monte Carlo Virtual Element (MLMC-VE) method for this purpose. VE methods allow, simultaneously, the natural construction of multilevel discretizations on general polytopal meshes required in the multilevel framework, and guarantee stability and approximation properties across levels. Given the virtual nature of the basis functions, the design and analysis of MLMC-VE is far from being a straightforward extension of the FEM case, as it requires careful use of the projectors of the virtual functions in the polynomial space.

We first derive novel VE a priori error estimates for linear quantities of interest of the solution to deterministic elliptic problems. Building on this result, we present Monte Carlo Virtual Element (MC-VE) estimators for the expectation of linear quantities of interest of the solution to stochastic elliptic problems with a random diffusion coefficient. We prove error bounds for the MC-VE estimator, showing that as expected, the MC convergence rate is  $1/2$ , so to halve the error, one should quadruple the sample size  $M$ . Given that, for each sample, the MC-VE method requires the solution of a VE problem, the overall computational complexity of the MC-VE method might become very large even for not too small error tolerances, as in the classical Monte Carlo setting. Next, to reduce the computational complexity of the MC-VE method, we propose a multilevel Monte Carlo Virtual Element (MLMC-VE) that uses a hierarchy of VE discretizations. The MLMC-VE estimator significantly reduces the sample count on the finest meshes, thereby improving efficiency. Exploiting the linearity of the expectation, the method computes the difference between expectations of the solutions (or their QoI) at two consecutive levels of accuracy. The advantage lies in reducing problem variance, thereby decreasing the number of samples required at finer (and more expensive) levels. We prove convergence, discuss computational complexity, and show that it significantly decreases the computational cost to achieve a given accuracy compared to standard Monte Carlo methods. Numerical experiments confirm the theoretical error estimates and show the computational benefits of our approach. To the best of our knowledge, this is one of the first results that incorporate stochastic modeling techniques into the VE framework to address uncertainty and to develop stochastic virtual element methods, in which deterministic VE spaces are coupled with stochastic approximation spaces to represent random variables and stochastic fields. In this direction, we also mention the recent paper [23], which proposes stochastic VE methods for two- and three-dimensional linear elasticity problems.

The rest of the paper is organized as follows. In Section 2, we introduce the model problem, namely an elliptic PDE with stochastic diffusion coefficient, together with the functional setting. In Section 3, we provide a brief overview of the conforming Virtual Element discretization of the corresponding deterministic model problem, and in Section 4, we derive novel *a priori* error estimates for a linear QoI. In Section 5, we introduce the MC-VE method and derive estimates for the statistical approximation error for both the solution and the linear quantity of interest. The extension to the multilevel setting is addressed in Section 6, where we introduce and establish error estimates for the MLMC-VE estimators, for both the solution and the quantity of interest. Extensive numerical experiments are presented and discussed in Section 7. Finally, in Section 8, we draw some conclusions and discuss possible future extensions.

## 2 Problem setting

In this section, we introduce the model problem considered in this work, namely, an elliptic partial differential equation with a stochastic diffusion coefficient.

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, where  $\Omega$  denotes the set of outcomes,  $\mathcal{F}$  the  $\sigma$ -algebra of events, and  $\mathbb{P}: \Omega \rightarrow [0, 1]$  a probability measure. Moreover, let  $D \subset \mathbb{R}^2$  be an open, bounded, convex polygonal domain with boundary  $\partial D$ . In this work, we focus on the following elliptic PDE with stochastic diffusion coefficient: given  $f \in L^2(D)$  and  $a: \Omega \times D \rightarrow \mathbb{R}$ , find  $u: \Omega \times D \rightarrow \mathbb{R}$  such that, almost surely in  $\Omega$ ,

$$\begin{cases} -\nabla \cdot (a \nabla u) = f & \text{in } D, \\ u = 0 & \text{on } \partial D. \end{cases} \quad (1)$$

Given the realization  $a(\omega, \cdot)$ , the weak formulation of (1) reads: find  $u(\omega, \cdot) \in V = H_0^1(D)$  such that, for all

$v \in V$ ,

$$\int_D a(\omega, \cdot) \nabla u(\omega, \cdot) \cdot \nabla v \, dx = \int_D f v \, dx, \quad (2)$$

which is well-posed by the Lax-Milgram lemma, provided that  $a(\omega, \cdot)$  is uniformly bounded below and above in  $D$  by positive constants. We make the following assumption.

**Assumption 2.1.** *For almost all  $\omega \in \Omega$ , the realizations  $a(\omega, \cdot)$  of the diffusion coefficient belong to  $L^\infty(D)$  and satisfy*

$$0 < a_{\min}(\omega) \leq a(\omega, x) \leq a_{\max}(\omega) < \infty \quad \text{a.e. in } D,$$

where

$$a_{\min}(\omega) = \text{ess inf}_{x \in D} a(\omega, x), \quad a_{\max}(\omega) = \text{ess sup}_{x \in D} a(\omega, x).$$

Note, in particular, that a random field  $a(\omega, x)$  satisfying Assumption 2.1 may not be uniformly bounded over  $\Omega$ . Under Assumption 2.1, repeated application of the Lax-Milgram lemma yields the almost surely well-posedness of problem (1), namely, for almost every  $\omega \in \Omega$ , problem (2) admits a unique solution  $u(\omega, \cdot) \in V$  and there exists a positive constant  $C$  independent of  $\omega$  such that

$$\|u(\omega, \cdot)\|_V \leq C a_{\min}(\omega)^{-1} \|f\|_{L^2(D)}.$$

**Assumption 2.2.** *For any  $p \geq 1$ ,  $a_{\min}^{-1} \in L^p(\Omega)$  and  $a \in L^p(\Omega, C^0(\bar{D}))$ .*

Under Assumption 2.2, it can be proved that  $u \in L^p(\Omega, V)$  for any  $p \geq 1$  and, for a constant  $C > 0$ ,

$$\|u\|_{L^p(\Omega, V)} \leq C \|a_{\min}^{-1}\|_{L^p(\Omega)} \|f\|_{L^2(D)}, \quad (3)$$

where the Bochner norm is defined by

$$\|\cdot\|_{L^p(\Omega, V)} = \left( \int_\Omega |\cdot|_{H^1(D)}^p \right)^{1/p}.$$

Note that Assumption 2.2 holds, for example, for random fields  $a(\omega, x)$  that are uniformly bounded in  $\Omega \times D$ , as well as for log-normal random fields  $a(\omega, x) = e^{Y(\omega, x)}$  with  $Y: \Omega \times \mathbb{R}$  mean-zero Gaussian random field fulfilling, for some  $L, s > 0$  and for all  $x_1, x_2 \in \bar{D}$ ,

$$\mathbb{E}[|Y(\omega, x_1) - Y(\omega, x_2)|^2] \leq L \|x_1 - x_2\|_2^s.$$

**Remark 2.1.** *For simplicity, we have assumed that the forcing term is deterministic. Nevertheless, the results mentioned above, as well as the forthcoming analysis, extend to the more general case where  $f = f(\omega, x)$  is a sufficiently regular random field (namely,  $f \in L^p(\Omega, L^2(D))$  for any  $p > 1$ ). More general boundary conditions, other than homogeneous Dirichlet, can be handled with minor modifications. Moreover, the proposed methodology naturally extends to the three-dimensional case  $D \subset \mathbb{R}^3$ . ■*

The present work aims to approximate the expected value of the solution  $u$  to problem (1), i.e.,

$$\mathbb{E}[u] = \int_\Omega u(\omega, \cdot) \, d\mathbb{P}(\omega) \in V$$

and the expected value of a quantity of interest  $\mathcal{Q}: V \rightarrow \mathbb{R}$  of  $u$ , i.e.,

$$\mathbb{E}[\mathcal{Q}(u)] = \int_\Omega \mathcal{Q}(u(\omega, \cdot)) \, d\mathbb{P}(\omega) \in \mathbb{R},$$

where we assume that  $\mathcal{Q}: V \rightarrow \mathbb{R}$  is a linear operator defined, for a given  $q \in L^2(D)$ , as

$$\mathcal{Q}(v) = \int_D qv. \quad (4)$$

### 3 The Virtual Element Method

In this section, we provide a brief overview of the conforming virtual element (VE) method and the main theoretical results that form the basis for the design and study of the Monte Carlo and Multilevel Monte Carlo VE discretizations introduced in this paper. Here and in the following, we will make use of the notation  $\lesssim$  to denote a bound that holds up to a positive multiplicative constant  $C$  that may vary at different occurrences, but it is independent of the mesh size, the polynomial degree, and the model's parameters.

We consider the following elliptic PDE with a *deterministic* diffusion coefficient: given  $f: D \rightarrow \mathbb{R}$  and  $\alpha: D \rightarrow \mathbb{R}$ , find  $u: D \rightarrow \mathbb{R}$  such that

$$\begin{cases} -\nabla \cdot (\alpha \nabla u) = f & \text{in } D, \\ u = 0 & \text{on } \partial D, \end{cases} \quad (5)$$

where we assume  $f \in L^2(D)$  and  $\alpha$  uniformly bounded from above and below in  $D$  by positive constants being

$$\alpha_{min} = \text{ess inf}_{x \in D} \alpha(x), \quad \alpha_{max} = \text{ess sup}_{x \in D} \alpha(x). \quad (6)$$

The weak formulation of (5) reads as follows: find  $u \in H_0^1(D)$  such that

$$A(u, v) = F(v) \quad \forall v \in H_0^1(D), \quad (7)$$

where

$$A(u, v) : H_0^1(D) \times H_0^1(D) \rightarrow \mathbb{R}, \quad A(u, v) = \int_D \alpha \nabla u \cdot \nabla v, \quad (8)$$

$$F(v) : H_0^1(D) \rightarrow \mathbb{R}, \quad F(v) = \int_D f v. \quad (9)$$

The remainder of this section is devoted to introduce the VE spaces, the bilinear forms, and the corresponding discrete problem.

#### 3.1 Virtual Element spaces

Let  $\mathcal{T}_h$  be a decomposition of  $D$  into  $n_P$  non-overlapping (open) polygons  $E_\ell$  with flat faces, i.e.,  $\bar{D} = \cup_{\ell=1}^{n_P} \bar{E}_\ell$  with  $E_\ell \cap E_{\ell'} = \emptyset$  for  $\ell \neq \ell'$ . Let  $h_E = \text{diam}(E)$  denote the diameter of the element  $E$ , and  $h = \max_{E \in \mathcal{T}_h} h_E$  denote the meshsize. We assume that there exist constants  $\gamma, c > 0$  independent of  $h$  and  $E$  such that

- each element  $E \in \mathcal{T}_h$  is star-shaped with respect to a ball of radius  $\gamma h_E$ ,
- the distance between any two vertices of  $E$  is larger than  $c h_E$ .

We denote by  $H^1(\mathcal{T}_h)$  the broken  $H^1$  space on  $\mathcal{T}_h$ :

$$H^1(\mathcal{T}_h) = \{v \in L^2(D) : v|_E \in H^1(E) \quad \forall E \in \mathcal{T}_h\}$$

and by  $|\cdot|_{H^1(\mathcal{T}_h)}^2 = \sum_{E \in \mathcal{T}_h} |\cdot|_{H^1(E)}^2$  the broken  $H^1$  seminorm. Let  $p \in \mathbb{N}$  be fixed. For any  $E \in \mathcal{T}_h$ , we introduce the following notation:

- $\mathbb{P}_p(E)$  is the set of polynomials on  $E$  of total degree less than or equal to  $p$ ;
- $\mathbb{B}(\partial E) = \{v \in C^0(\partial E) \text{ s.t. } v|_e \in \mathbb{P}_p(e) \text{ for all edges } e \subset \partial E\}$ ;
- $\Pi_p^{\nabla, E} : H^1(E) \rightarrow \mathbb{P}_p(E)$  is the energy projection operator, which is defined, up to a constant, by

$$\int_E \nabla(w - \Pi_p^{\nabla, E} w) \cdot \nabla q_p \, dx = 0 \quad \forall q_p \in \mathbb{P}_p(E). \quad (10)$$

To fix the constant and complete the definition of  $\Pi_p^{\nabla, E} w$ , we further require

$$\begin{cases} \frac{1}{|\partial E|} \int_{\partial E} (w - \Pi_p^{\nabla, E} w) \, ds = 0, & \text{if } p = 1, \\ \frac{1}{|E|} \int_E (w - \Pi_p^{\nabla, E} w) \, dx = 0, & \text{if } p > 1; \end{cases}$$

- $\Pi_p^{0, E} : L^2(E) \rightarrow \mathbb{P}_p(E)$  is the  $L^2$ -orthogonal projection operator defined by

$$(q_p, w - \Pi_p^{0, E} w)_{L^2(E)} = 0 \quad \forall q_p \in \mathbb{P}_p(E). \quad (11)$$

Since  $\Pi_p^{\nabla,E}$  and  $\Pi_p^{0,E}$  provide the best approximation in  $\mathbb{P}_p(E)$  with respect to the  $H^1(E)$  seminorm and in the  $L^2(E)$  norm, respectively, from [7, Lemma 4.2], it follows that, for all  $u \in H^{s+1}(E)$ ,  $s \geq 1$ , letting  $\mu_s = \min\{s, p\}$ ,

$$|u - \Pi_p^{\nabla,E} u|_{H^1(E)} \lesssim \frac{h_E^{\mu_s}}{p^s} \|u\|_{H^{s+1}(E)}, \quad (12)$$

$$\|u - \Pi_p^{0,E} u\|_{L^2(E)} \lesssim \frac{h_E^{\mu_s+1}}{p^{s+1}} \|u\|_{H^{s+1}(E)}. \quad (13)$$

Moreover, the global operators  $\Pi_{h,p}^{\nabla}, \Pi_{h,p}^0$  are defined as  $\Pi_{h,p}^{(\cdot)}|_E = \Pi_{h,p}^{(\cdot),E}$  with  $(\cdot) = \nabla, 0$  and, for all  $u \in H^{s+1}(D)$ ,  $s \geq 1$ , fulfill

$$|u - \Pi_{h,p}^{\nabla} u|_{H^1(\mathcal{T}_h)} \lesssim \frac{h^{\mu_s}}{p^s} \|u\|_{H^{s+1}(D)}, \quad (14)$$

$$\|u - \Pi_{h,p}^0 u\|_{L^2(D)} \lesssim \frac{h^{\mu_s+1}}{p^{s+1}} \|u\|_{H^{s+1}(D)}. \quad (15)$$

Additionally, for all  $u \in H^1(D)$ , we have

$$\|u - \Pi_{h,p}^0 u\|_{L^2(D)} \lesssim \frac{h}{p} |u|_{H^1(D)}. \quad (16)$$

Finally, from (10), we have the following local and global stability properties:

$$|\Pi_p^{\nabla,E} u|_{H^1(E)} \leq |u|_{H^1(E)}, \quad |\Pi_{h,p}^{\nabla} u|_{H^1(\mathcal{T}_h)} \leq |u|_{H^1(D)}. \quad (17)$$

We can now introduce the local enhanced VE space of order  $p \geq 1$ :

$$W_p^E = \left\{ w \in V_p^E \text{ s.t. } (w - \Pi_p^{\nabla,E} w, q)_{L^2(E)} = 0 \text{ for all } q \in \mathbb{P}_p(E)/\mathbb{P}_{p-2}(E) \right\}, \quad (18)$$

where  $V_p^E$  denotes the local augmented VE space

$$V_p^E = \{w \in H^1(E) \text{ s.t. } w \in \mathbb{B}_p(\partial E) \text{ and } \Delta w \in \mathbb{P}_p(E)\},$$

and  $\mathbb{P}_p(E)/\mathbb{P}_{p-2}(E)$  denotes the set of polynomials of total degree  $p$  on  $E$  that are  $L^2$ -orthogonal to all polynomials of total degree  $p-2$  on  $E$  (with the convention  $\mathbb{P}_{-1} = \emptyset$ ). Note, in particular, that  $\mathbb{P}_p(E) \subset W_{h,p}^E(E)$ . The space  $W_{h,p}^E$  is equipped with the following set of (local) degrees of freedom (DOFs):

- nodal values at all  $n_E$  vertices of the polygon  $E$ ;
- nodal values at the  $p-1$  internal Gauss-Lobatto quadrature points of each edge  $e \in \partial E$ ;
- (for  $p \geq 2$ ) moments up to order  $p-2$  in  $E$ , namely, for  $w \in W_p^E$ ,

$$(w, q_{p-2})_{L^2(E)} \quad \text{for all } q_{p-2} \in \mathbb{P}_{p-2}.$$

We have that  $\dim(W_p^E) = n_E p + \frac{p(p-1)}{2}$ . It is important to note that both the energy and the  $L^2$ -orthogonal projection operators are computable solely in terms of the DOFs defined above. The global enhanced VE space of degree  $p$  is defined by

$$W_{h,p} = \{v \in H_0^1(D) \text{ s.t. } v|_E \in W_p^E \text{ for all } E \in \mathcal{T}_h\}. \quad (19)$$

It is equipped with the following set of (global) DOFs:

- nodal values at all  $n_V$  vertices of  $\mathcal{T}_h$ ;
- nodal values at the  $p-1$  internal Gauss-Lobatto quadrature points of each of the  $n_e$  edges of  $\mathcal{T}_h$ ;
- (for  $p \geq 2$ ) moments up to order  $p-2$  in each of the  $n_P$  polygons of  $\mathcal{T}_h$ , namely, for  $w \in W_{h,p}$ ,

$$(w, q_{p-2})_{L^2(E)} \quad \text{for all } q_{p-2} \in \mathbb{P}_{p-2}, \text{ for all } E \in \mathcal{T}_h,$$

and it has dimension  $N_{h,p} = \dim(W_{h,p}) = n_V + (p-1)n_e + n_P \frac{p(p-1)}{2}$ . It was shown in [22, Corollary 2.3] that, for any  $w \in H^{s+1}(D)$ ,  $s \geq 1$ , there exists a VE function  $w_I \in W_{h,p}$  such that

$$|w - w_I|_{H^1(D)} \lesssim \frac{h^{\mu_s}}{p^{s-1}} \|w\|_{H^{s+1}(D)}, \quad (20)$$

where we recall the definition  $\mu_s = \min\{s, p\}$ .

**Remark 3.1.** We point out that, although the above result (20) is expected to be valid with  $p^s$  instead of  $p^{s-1}$ , in the following, we use (20), which is the current state of the art. ■

### 3.2 Virtual Element bilinear forms

Let  $A^E(u, v) = \int_E \alpha(x) \nabla u \cdot \nabla v \, dx$ . Given an arbitrary pair of VE functions  $v_{h,p}, w_{h,p} \in W_p^E$ , the quantity  $A^E(v_{h,p}, w_{h,p}) = \int_E \alpha \nabla v_{h,p} \cdot \nabla w_{h,p}$  is not computable in terms of the VE DOFs. Thus, we introduce the computable approximation  $A_h^E: W_p^E \times W_p^E \rightarrow \mathbb{R}$ , given by

$$A_h^E(v_{h,p}, w_{h,p}) = A^E(\Pi_p^{\nabla, E} v_{h,p}, \Pi_p^{\nabla, E} w_{h,p}) + S^E((Id - \Pi_p^{\nabla, E})v_h, (Id - \Pi_p^{\nabla, E})w_h), \quad (21)$$

where  $S^E: W_p^E \times W_p^E \rightarrow \mathbb{R}$  denotes a computable, symmetric, stabilizing bilinear form satisfying, for all  $w_h \in W_p^E$  with  $\Pi_p^{\nabla, E} w_{h,p} = 0$ ,

$$s_*(p)A^E(w_{h,p}, w_{h,p}) \lesssim S^E(w_{h,p}, w_{h,p}) \lesssim s^*(p)A^E(w_{h,p}, w_{h,p}) \quad (22)$$

for some positive constants  $s_*(p), s^*(p)$ , with  $s^*(p) \geq 1$ , which may depend only on  $p$ .

**Remark 3.2.** *In our numerical experiments, we define the stabilizing form  $S^E$  as follows:*

$$S^E(v_{h,p}, w_{h,p}) = \alpha_E \sum_{r=1}^{N_E} \text{DOF}_r \left( (Id - \Pi_p^{\nabla, E})v_{h,p} \right) \text{DOF}_r \left( (Id - \Pi_p^{\nabla, E})w_{h,p} \right),$$

where  $\alpha_E = \frac{1}{|E|} \int_E \alpha \, dx, N_E = \dim(W_{h,p}^E)$ , and  $\{\text{DOF}_r\}_{r=1}^{N_E}$  denotes the set of local DOFs introduced in Section 3.1. For  $\alpha \in W^{1,\infty}(D)$  and sufficiently fine meshes,  $S^E$  satisfies (22).  $\blacksquare$

The global VE bilinear form  $A_h: W_{h,p} \times W_{h,p} \rightarrow \mathbb{R}$  is then defined, for all  $v_{h,p}, w_{h,p} \in W_{h,p}$ , as

$$A_h(v_{h,p}, w_{h,p}) = \sum_{E \in \mathcal{T}_h} A_h^E(v_{h,p}, w_{h,p}).$$

The global VE bilinear form is continuous and coercive: for all  $v_{h,p}, w_{h,p} \in W_{h,p}$ ,

$$\begin{aligned} A_h(v_{h,p}, w_{h,p}) &\lesssim s^*(p) \alpha_{max} |v_{h,p}|_{H^1(D)} |w_{h,p}|_{H^1(D)}, \\ A_h(v_{h,p}, v_{h,p}) &\gtrsim s_*(p) \alpha_{min} |v_{h,p}|_{H^1(D)}^2. \end{aligned}$$

### 3.3 The discrete problem

The VE discretization of problem (5) reads as follows: find  $u_{h,p} \in W_{h,p}$  such that

$$A_h(u_{h,p}, v_{h,p}) = F_h(v_{h,p}) \quad \forall v_{h,p} \in W_{h,p}, \quad (23)$$

where

$$F_h(v_{h,p}) = \sum_{E \in \mathcal{T}_h} \int_E \left( \Pi_p^{0,E} f \right) v_{h,p} = \sum_{E \in \mathcal{T}_h} \int_E \left( \Pi_p^{0,E} f \right) \left( \Pi_p^{0,E} v_{h,p} \right).$$

From the stability of the  $L^2$  projection in the  $L^2$  norm and the Poincaré inequality, for all  $v_{h,p} \in W_{h,p}$ , we have

$$|F_h(v_{h,p})| \lesssim \|f\|_{L^2(D)} |v_{h,p}|_{H^1(D)},$$

which, together with the coercivity of  $A_h$ , implies

$$|u_{h,p}|_{H^1(D)} \lesssim s_*(p)^{-1} \alpha_{min}^{-1} \|f\|_{L^2(D)}, \quad (24)$$

Then, the discrete problem (23) is well-posed. Moreover, if  $u \in H^{s+1}(D)$ ,  $s \geq 1$ , the following error estimates has been proven, e.g., in [7, Theorem 4.1, Remark 4.4, and Remark 4.5]:

$$\begin{aligned} |u - u_{h,p}|_{H^1(D)} &\lesssim \frac{s^*(p)}{s_*(p)} \frac{\alpha_{max}}{\alpha_{min}} \frac{h^{\mu_s}}{p^s} \|u\|_{H^{s+1}(D)}, \\ \|u - u_{h,p}\|_{L^2(D)} &\lesssim \frac{s^*(p)}{s_*(p)} \frac{\alpha_{max}}{\alpha_{min}} \frac{h^{\mu_s+1}}{p^{s+1}} \|u\|_{H^{s+1}(D)}, \end{aligned} \quad (25)$$

with  $\mu_s = \min\{s, p\}$ .

## 4 A priori Virtual Element error estimates for a linear quantity of interest

In this section, we present novel *a priori* VE error estimates for a linear QoI of the solution in the deterministic setting.

Recalling that  $V = H_0^1(D)$ , and  $\mathcal{Q} : V \rightarrow \mathbb{R}$  is the linear operator defined in (4), we are interested in estimating  $\mathcal{Q}(u) - \mathcal{Q}(u_{h,p})$ , where  $u \in V$  is the solution of (7), and  $u_{h,p} \in W_{h,p}$  is the VE approximation to  $u$ , i.e., it satisfies (23). For  $v_{h,p} \in W_{h,p}$ , let us define

$$\mathcal{Q}_h(v_{h,p}) = \int_D (\Pi_{p,h}^0 q) v_{h,p}, \quad (26)$$

which, unlike  $\mathcal{Q}(u_{h,p})$ , is a computable quantity, since it holds

$$\mathcal{Q}_h(v_{h,p}) = \int_D (\Pi_{p,h}^0 q) (\Pi_{p,h}^0 v_{h,p}) = \mathcal{Q}(\Pi_{p,h}^0 v_{h,p}). \quad (27)$$

The following result is an extension of  $L^2$ -error estimates for virtual elements, cf. [12, Section 3.4].

**Theorem 4.1.** *Let  $z$  be the solution of the auxiliary problem*

$$-\nabla \cdot (\alpha \nabla z) = q \quad \text{in } D, \quad z = 0 \quad \text{on } \partial D. \quad (28)$$

Assume that  $f \in H^{s-1}(D)$ ,  $q \in H^{t-1}(D)$ ,  $u \in H^{s+1}(D)$ ,  $z \in H^{t+1}(D)$ ,  $s, t \geq 1$ , and

$$\|z\|_{H^{t+1}(D)} \lesssim \alpha_{\min}^{-1} \|q\|_{H^{t-1}(D)}, \quad \|u\|_{H^{s+1}(D)} \lesssim \alpha_{\min}^{-1} \|f\|_{H^{s-1}(D)}. \quad (29)$$

Then, setting  $\mu = \min\{s, t, p\}$ , we have

$$|\mathcal{Q}(u - u_{h,p})| \lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{\alpha_{\max}}{\alpha_{\min}} \right)^2 \alpha_{\min}^{-1} \frac{h^{2\mu}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right), \quad (30)$$

$$|\mathcal{Q}(u) - \mathcal{Q}_h(u_{h,p})| \lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{\alpha_{\max}}{\alpha_{\min}} \right)^2 \alpha_{\min}^{-1} \frac{h^{2\mu}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right). \quad (31)$$

*Proof.* Let us set  $\mu_s = \min\{s, p\}$  and  $\mu_t = \min\{t, p\}$ . We denote by  $z_I$  the VE function in  $W_{h,p}$  corresponding to  $z \in H^{t+1}(D)$  as in (20), for which there holds  $|z - z_I|_{H^1(D)} \lesssim \frac{h^{\mu_t}}{p^{t-1}} \|z\|_{H^{t+1}(D)}$ . Using the auxiliary problem (28), we have

$$|\mathcal{Q}(u - u_{h,p})| = |A(z, u - u_{h,p})| \leq |A(z - z_I, u - u_{h,p})| + |A(z_I, u - u_{h,p})| \quad (32)$$

$$=: (I) + (II). \quad (33)$$

By the continuity of  $A$ , (25), (20), and (29), there holds

$$\begin{aligned} (I) &\leq \alpha_{\max} |u - u_{h,p}|_{H^1(D)} |z - z_I|_{H^1(D)} \lesssim \frac{s^*(p)}{s_*(p)} \frac{\alpha_{\max}^2}{\alpha_{\min}} \frac{h^{\mu_s + \mu_t}}{p^{s+t-1}} \|u\|_{H^{s+1}(D)} \|z\|_{H^{t+1}(D)} \\ &\lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{\alpha_{\max}}{\alpha_{\min}} \right)^2 \alpha_{\min}^{-1} \frac{h^{\mu_s + \mu_t}}{p^{s+t-1}} \|f\|_{H^{s-1}(D)} \|q\|_{H^{t-1}(D)}. \end{aligned}$$

For the second term, we write

$$(II) \leq \left| \sum_E A^E(z_I - \Pi_p^{\nabla, E} z_I, u - u_{h,p}) \right| + \left| \sum_E A^E(\Pi_p^{\nabla, E} z_I, u - u_{h,p}) \right| =: (III) + (IV),$$

and estimate (III) and (IV) separately. Before doing so, we prove that

$$\sum_E |z_I - \Pi_p^{\nabla, E} z_I|_{H^1(E)}^2 \lesssim \frac{h^{2\mu_t}}{p^{2t-2}} \|z\|_{H^{t+1}(D)}^2, \quad (34)$$

$$\sum_E |u_{h,p} - \Pi_p^{\nabla, E} u_{h,p}|_{H^1(E)}^2 \lesssim \frac{h^{2\mu_s}}{p^{2s}} \|u\|_{H^{s+1}(D)}^2 \quad (35)$$

To prove (34), we use the triangle inequality, the stability of  $\Pi_p^{\nabla,E}$  in (17), and the estimates (20) and (14):

$$\begin{aligned}
\sum_E |z_I - \Pi_p^{\nabla,E} z_I|_{H^1(E)}^2 &\leq \sum_E \left( |z_I - z|_{H^1(E)} + |z - \Pi_p^{\nabla,E} z|_{H^1(E)} + |\Pi_p^{\nabla,E}(z - z_I)|_{H^1(E)} \right)^2 \\
&\lesssim \sum_E \left( |z_I - z|_{H^1(E)} + |z - \Pi_p^{\nabla,E} z|_{H^1(E)} \right)^2 \\
&\lesssim \left( |z_I - z|_{H^1(D)}^2 + |z - \Pi_p^{\nabla,E} z|_{H^1(\mathcal{T}_h)}^2 \right) \\
&\lesssim \frac{h^{2\mu_t}}{p^{2t-2}} \|z\|_{H^{t+1}(D)}^2.
\end{aligned}$$

The proof of (35) follows the same lines, replacing  $z_I$  by  $u_{h,p}$ ,  $z$  by  $u$ , and using (25) instead of (20). We now estimate (III). Using the continuity of  $A^E$ , the discrete Cauchy-Schwarz inequality, and the estimates (34) and (25), we obtain

$$\begin{aligned}
(III) &\leq \alpha_{max} \sum_E |z_I - \Pi_p^{\nabla,E} z_I|_{H^1(E)} |u - u_{h,p}|_{H^1(E)} \\
&\leq \alpha_{max} \left( \sum_E |z_I - \Pi_p^{\nabla,E} z_I|_{H^1(E)}^2 \right)^{1/2} |u - u_{h,p}|_{H^1(D)} \\
&\lesssim \frac{s^*(p)}{s_*(p)} \frac{\alpha_{max}^2 h^{\mu_s + \mu_t}}{\alpha_{min} p^{s+t-1}} \|z\|_{H^{t+1}(D)} \|u\|_{H^{s+1}(D)} \\
&\lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{\alpha_{max}}{\alpha_{min}} \right)^2 \alpha_{min}^{-1} \frac{h^{\mu_s + \mu_t}}{p^{s+t-1}} \|q\|_{H^{t-1}(D)} \|f\|_{H^{s-1}(D)},
\end{aligned}$$

where, in the last step, we used (29). In order to estimate (IV), let us preliminary observe that there holds

$$\begin{aligned}
A^E(z_I, \Pi_p^{\nabla,E} u_{h,p}) &= A^E(\Pi_p^{\nabla,E} z_I, \Pi_p^{\nabla,E} u_{h,p}) \\
&= A_h^E(z_I, u_{h,p}) - S^E((I - \Pi_p^{\nabla,E}) z_I, (I - \Pi_p^{\nabla,E}) u_{h,p}),
\end{aligned}$$

where we employed the orthogonality property of  $\Pi_p^{\nabla,E}$  and the definition of  $A_h^E$ . Hence, we have

$$\begin{aligned}
(IV) &= \left| \sum_E A^E(z_I, \Pi_p^{\nabla,E}(u - u_{h,p})) \right| \\
&= \left| \sum_E (A^E(z_I, \Pi_p^{\nabla,E} u) - A_h^E(z_I, u_{h,p}) + S^E((I - \Pi_p^{\nabla,E}) z_I, (I - \Pi_p^{\nabla,E}) u_{h,p})) \right| \\
&= \left| \sum_E (A^E(z_I, \Pi_p^{\nabla,E} u - u) + S^E((I - \Pi_p^{\nabla,E}) z_I, (I - \Pi_p^{\nabla,E}) u_{h,p})) \right. \\
&\quad \left. + F(z_I) - F_h(z_I) \right| \\
&\leq \left| \sum_E A^E(z_I - \Pi_p^{\nabla,E} z_I, \Pi_p^{\nabla,E} u - u) \right| + \left| \sum_E S^E((I - \Pi_p^{\nabla,E}) z_I, (I - \Pi_p^{\nabla,E}) u_{h,p}) \right| \\
&\quad + |F(z_I) - F_h(z_I)|.
\end{aligned}$$

Since

$$\begin{aligned}
A^E(z_I - \Pi_p^{\nabla,E} z_I, \Pi_p^{\nabla,E} u - u) &\leq \alpha_{max} |z_I - \Pi_p^{\nabla,E} z_I|_{H^1(E)} |\Pi_p^{\nabla,E} u - u|_{H^1(E)} \\
&\lesssim \alpha_{max} |z_I - \Pi_p^{\nabla,E} z_I|_{H^1(E)}^2 + \alpha_{max} |\Pi_p^{\nabla,E} u - u|_{H^1(E)}^2,
\end{aligned}$$

using (34) and the estimate in (14), we obtain

$$\begin{aligned}
&\left| \sum_E A^E(z_I - \Pi_p^{\nabla,E} z_I, \Pi_p^{\nabla,E} u - u) \right| \\
&\lesssim \alpha_{max} \frac{h^{2\mu_t}}{p^{2t-2}} \|z\|_{H^{t+1}(D)}^2 + \alpha_{max} \frac{h^{2\mu_s}}{p^{2s}} \|u\|_{H^{s+1}(D)}^2 \\
&\lesssim \frac{\alpha_{max}}{\alpha_{min}} \alpha_{min}^{-1} \frac{h^{2 \min\{\mu_s, \mu_t\}}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right),
\end{aligned}$$

where, in the last step, we used again (29). Furthermore, the bilinearity and symmetry of  $S^E$ , together with the equivalence (22), imply

$$\begin{aligned}
& S^E((I - \Pi_p^{\nabla,E})z_I, (I - \Pi_p^{\nabla,E})u_{h,p}) \\
&= \frac{1}{2}S^E((I - \Pi_p^{\nabla,E})z_I + (I - \Pi_p^{\nabla,E})u_{h,p}, (I - \Pi_p^{\nabla,E})z_I + (I - \Pi_p^{\nabla,E})u_{h,p}) \\
&\quad - \frac{1}{2}S^E((I - \Pi_p^{\nabla,E})z_I, (I - \Pi_p^{\nabla,E})z_I) - \frac{1}{2}S^E((I - \Pi_p^{\nabla,E})u_{h,p}, (I - \Pi_p^{\nabla,E})u_{h,p}) \\
&\lesssim s^*(p)A^E((I - \Pi_p^{\nabla,E})z_I + (I - \Pi_p^{\nabla,E})u_{h,p}, (I - \Pi_p^{\nabla,E})z_I + (I - \Pi_p^{\nabla,E})u_{h,p}) \\
&\leq s^*(p)\alpha_{max} |(I - \Pi_p^{\nabla,E})z_I + (I - \Pi_p^{\nabla,E})u_{h,p}|_{H^1(E)}^2 \\
&\lesssim s^*(p)\alpha_{max} |(I - \Pi_p^{\nabla,E})z_I|_{H^1(E)}^2 + s^*(p)\alpha_{max} |(I - \Pi_p^{\nabla,E})u_{h,p}|_{H^1(E)}^2.
\end{aligned}$$

Therefore, using (34) and (35), we obtain

$$\begin{aligned}
& \left| \sum_E S^E((I - \Pi_p^{\nabla,E})z_I, (I - \Pi_p^{\nabla,E})u_{h,p}) \right| \\
&\lesssim s^*(p)\alpha_{max} \frac{h^{2 \min\{\mu_s, \mu_t\}}}{p^{2 \min\{s, t-1\}}} \left( \|z\|_{H^{t+1}(D)}^2 + \|u\|_{H^{s+1}(D)}^2 \right) \\
&\lesssim s^*(p) \frac{\alpha_{max}}{\alpha_{min}} \alpha_{min}^{-1} \frac{h^{2 \min\{\mu_s, \mu_t\}}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right),
\end{aligned}$$

where, in the last step, we used again (29). Finally, for  $|F(z_I) - F_h(z_I)|$ , we have

$$\begin{aligned}
|F(z_I) - F_h(z_I)| &= |(F - F_h)(z_I - z) + (F - F_h)(z)| \\
&= \left| \sum_E \int_E (f - \Pi_p^{0,E} f)(z_I - z) + \sum_E \int_E (f - \Pi_p^{0,E} f)(z - \Pi_p^{0,E} z) \right| \\
&= \left| \sum_E \int_E (f - \Pi_p^{0,E} f)(I - \Pi_p^{0,E})(z_I - z) + \sum_E \int_E (f - \Pi_p^{0,E} f)(z - \Pi_p^{0,E} z) \right| \\
&\leq \|f - \Pi_{h,p}^0 f\|_{L^2(D)} (\|(I - \Pi_{h,p}^0)(z_I - z)\|_{L^2(D)} + \|z - \Pi_{h,p}^0 z\|_{L^2(D)}) \\
&\lesssim \|f - \Pi_{h,p}^0 f\|_{L^2(D)} \left( \frac{h}{p} |z - z_I|_{H^1(D)} + \|z - \Pi_p^0 z\|_{L^2(D)} \right) \\
&\lesssim \frac{h^{\min\{s, p+2\} + \mu_t}}{p^{s+t-1}} \|f\|_{H^{s-1}(D)} \|z\|_{H^{t+1}(D)} \\
&\lesssim \alpha_{min}^{-1} \frac{h^{\min\{s, p+2\} + \mu_t}}{p^{s+t-1}} \|f\|_{H^{s-1}(D)} \|q\|_{H^{t-1}(D)}, \tag{36}
\end{aligned}$$

where we also employed (16) and the error estimate  $\|f - \Pi_p^{0,E} f\|_{L^2(D)} \lesssim \frac{h^{\min\{s-1, p+1\}}}{p^{s-1}} \|f\|_{H^{s-1}(D)}$ , which holds as  $f \in H^{s-1}(D)$ . Collecting the above results, we obtain, for the term (IV),

$$(IV) \lesssim s^*(p) \frac{\alpha_{max}}{\alpha_{min}} \alpha_{min}^{-1} \frac{h^{2 \min\{\mu_s, \mu_t\}}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right).$$

This, together with the estimate obtained for (III), gives

$$(II) \lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{\alpha_{max}}{\alpha_{min}} \right)^2 \alpha_{min}^{-1} \frac{h^{2 \min\{\mu_s, \mu_t\}}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right),$$

which, combined with the estimate obtained for (I), yields

$$|\mathcal{Q}(u - u_{h,p})| \lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{\alpha_{max}}{\alpha_{min}} \right)^2 \alpha_{min}^{-1} \frac{h^{2 \min\{\mu_s, \mu_t\}}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right).$$

Then, estimate (30) follows, owing to  $\min\{\mu_s, \mu_t\} = \min\{s, t, p\}$ .

To obtain (31), we observe that there holds

$$\begin{aligned}
|\mathcal{Q}(u) - \mathcal{Q}_h(u_{h,p})| &= |\mathcal{Q}(u) - \mathcal{Q}_h(u) + \mathcal{Q}_h(u) - \mathcal{Q}_h(u_{h,p})| \\
&= \left| \int_D (q - \Pi_{p,h}^0 q)u + \int_D \Pi_{p,h}^0 q(u - u_{h,p}) \right| \\
&= \left| \int_D (q - \Pi_{p,h}^0 q)(u - \Pi_{p,h}^0 u) + \int_D (\Pi_{p,h}^0 q - q)(u - u_{h,p}) + \mathcal{Q}(u - u_{h,p}) \right| \\
&\leq \|q - \Pi_{p,h}^0 q\|_{L^2(D)} (\|u - \Pi_{p,h}^0 u\|_{L^2(D)} + \|u - u_{h,p}\|_{L^2(D)}) + \mathcal{Q}(u - u_{h,p}).
\end{aligned}$$

Then, using the  $L^2$  projection error estimates (notice that  $q \in H^{t-1}(D)$ ) implies

$$\|q - \Pi_p^{0,E} q\|_{L^2(D)} \lesssim \frac{h^{\min\{t-1, p+1\}}}{p^{t-1}} \|q\|_{H^{t-1}(D)}$$

and the estimate in (25), we have

$$\begin{aligned} |\mathcal{Q}(u) - \mathcal{Q}_h(u_h)| &\lesssim \frac{\alpha^*(p)}{s_*(p)} \frac{\alpha_{max}}{\alpha_{min}} \frac{h^{\mu_s + \min\{t, p+2\}}}{p^{s+t}} \|q\|_{H^{t-1}(D)} \|u\|_{H^{s+1}(D)} + \mathcal{Q}(u - u_{h,p}) \\ &\lesssim \frac{\alpha^*(p)}{s_*(p)} \frac{\alpha_{max}}{\alpha_{min}} \alpha_{min}^{-1} \frac{h^{\mu_s + \min\{t, p+2\}}}{p^{s+t}} \|q\|_{H^{t-1}(D)} \|f\|_{H^{s-1}(D)} + \mathcal{Q}(u - u_{h,p}). \end{aligned}$$

This, together with (30) gives (31).  $\square$

**Remark 4.1.** The presence of  $p^{2 \min\{s, t-1\}}$  in the denominators in estimates (30) and (31) of Theorem 4.1 is a consequence of the use, in the proof, of the best approximant  $z_I$  in the “enhanced” VE space  $W_{h,p}$ , which was proven in to satisfy (20). However, a closer inspection of the proof of Theorem 4.1 shows that only the degrees of freedom of  $z_I$  are actually used, together with the consistency of method (23) when tested with the function  $z_I$ . Therefore, if the right-hand side  $F_h(v_{h,p})$  of method (23) were defined as

$$F_h(v_{h,p}) = \sum_{E \in \mathcal{T}_h} \int_E \left( \Pi_{p-2}^{0,E} f \right) v_{h,p},$$

that is employing  $\Pi_{p-2}^{0,E}$  instead of  $\Pi_p^{0,E}$ , one could use the best approximant in the “classical” VE space (see [7] for its definition), say  $z_I^c$ , instead of  $z_I$ , without modifying the argument. In fact, since only the degrees of freedom of the approximant enter the analysis, there is no difference between choosing  $z_I$  or  $z_I^c$ . We stress that modifying the definition of the right-hand side  $F_h$  is necessary for this argument. Indeed, with the original definition of  $F_h$ , the required consistency property would no longer hold when testing with  $z_I^c$ . Consequently, the estimate of term (IV) in the proof of Theorem 4.1 could not be reproduced after replacing  $z_I$  by  $z_I^c$ . The advantage of this alternative approach is that, for  $z_I^c$ , the analogue of estimate (20) would involve  $p^{\min\{s, t\}}$  in the denominator, instead of  $p^{\min\{s, t-1\}}$ ; see [7, Lemma 4.3], resulting in the presence of  $p^{2 \min\{s, t\}}$  instead of  $p^{2 \min\{s, t-1\}}$  in the denominators in estimates (30) and (31).  $\blacksquare$

## 5 The MC-VE method

In this section, we propose a method for approximating the expected value of the solution  $u$  of the elliptic problem with stochastic coefficient (1), as well as the expected value of  $\mathcal{Q}(u)$ .

A widely used strategy to discretize integrals in probability is the Monte Carlo method, which relies on the following operator:

$$E_M[w] = \frac{1}{M} \sum_{i=1}^M w^{(i)}, \quad (37)$$

where  $w^{(i)} = w(\omega_i)$ ,  $i = 1, \dots, M$ , are  $M$  independent identically distributed (i.i.d.) samples of the stochastic function  $w(\omega)$ . Possible examples are  $w^{(i)} = u(\omega_i, \cdot)$  and  $w^{(i)} = \mathcal{Q}(u(\omega_i, \cdot))$ , where  $u(\omega_i, \cdot)$  is the unique solution of (2) with  $\omega = \omega_i$ . We recall from [5] that the following holds (note that  $\mathbb{E}[w]$  is deterministic, whereas  $E_M[w]$  is a random variable because its value depends on the considered samples):

$$\|\mathbb{E}[w] - E_M[w]\|_{L^2(\Omega, V)} \leq M^{-1/2} \|w\|_{L^2(\Omega, V)} \quad \forall w \in L^2(\Omega, V). \quad (38)$$

As the aim of the present work is to approximate the expected value of the solution  $u$  to problem (1) as well as its quantity of interest  $\mathcal{Q}(u)$ , we combine MC with the VE method of order  $p$  defined over a polytopal mesh of meshsize  $h$ . As a result, we get the following MC-VE estimators for  $\mathbb{E}[u]$  and  $\mathbb{E}[\mathcal{Q}(u)]$ :

$$E_M[\Pi_{h,p}^\nabla u_{h,p}] = \frac{1}{M} \sum_{i=1}^M \Pi_{h,p}^\nabla u_{h,p}^{(i)} \in L^2(\Omega, H^1(\mathcal{T}_h)), \quad (39)$$

$$E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})] = \frac{1}{M} \sum_{i=1}^M \mathcal{Q}(\Pi_{h,p}^0 u_{h,p}^{(i)}) \in L^2(\Omega), \quad (40)$$

where  $u_{h,p}^{(i)} = u_{h,p}(\omega_i, \cdot)$ ,  $i = 1, \dots, M$ , are  $M$  i.i.d. samples of the stochastic VE solution  $u_{h,p}(\omega, \cdot)$ , obtained by solving  $M$  VE problems of the form (23), each corresponding to a sample  $a(\omega_i, \cdot)$  of the stochastic diffusion coefficient. We define the Bochner seminorm

$$\|\cdot\|_{L^2(\Omega, H^1(\mathcal{T}_h))} = \left( \int_{\Omega} |\cdot|_{H^1(\mathcal{T}_h)}^2 \right)^{1/2},$$

which is a norm for functions with zero mean on each element of  $\mathcal{T}_h$ . In  $L^2(\Omega, V)$ ,  $\|\cdot\|_{L^2(\Omega, H^1(\mathcal{T}_h))} = \|\cdot\|_{L^2(\Omega, V)}$ . Since in all the cases when it is used below it is actually a norm, we have defined it with the  $\|\cdot\|$  symbol.

**Remark 5.1.** *Although the natural choice in (39) would be  $E_M[u_{h,p}]$  (and  $E_M[\mathcal{Q}(u_{h,p})]$  in (40)), this is not practical, since the virtual function  $u_{h,p}$  is not known in closed form. For this reason, we use the projection  $\Pi_{h,p}^{\nabla} u_{h,p}$ , which is computable from the DOFs.  $\blacksquare$*

In the following proposition, we derive estimates for the statistical approximation error for both the solution and the quantity of interest. For the latter, we use the analogue of the estimate derived in Theorem 4.1 for a deterministic diffusion coefficient. In the case of problem (2), this estimate becomes, with  $\mu = \min\{s, t, p\}$ ,

$$\begin{aligned} & |\mathcal{Q}(u) - \mathcal{Q}_h(u_{h,p})| \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left( \frac{a_{max}(\omega)}{a_{min}(\omega)} \right)^2 a_{min}(\omega)^{-1} \frac{h^{2\mu}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right). \end{aligned} \quad (41)$$

**Proposition 5.1.** *Assume that  $u \in L^2(\Omega, H^{s+1}(D))$ ,  $s \geq 1$ . Then, the following error bound on  $\mathbb{E}[u]$  holds, with  $\mu_s = \min\{s, p\}$ :*

$$\begin{aligned} & \|\mathbb{E}[u] - E_M[\Pi_{h,p}^{\nabla} u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))} \\ & \lesssim \frac{s^*(p)}{s_*(p)} \max \left\{ \left\| \frac{a_{max}}{a_{min}} \right\|_{L^2(\Omega)}, \|a_{min}^{-1}\|_{L^2(\Omega)} \right\} \left( \frac{h^{\mu_s}}{p^s} + M^{-1/2} \right) \left( \|u\|_{L^2(\Omega, H^{s+1}(D))} + \|f\|_{L^2(D)} \right), \end{aligned} \quad (42)$$

where  $\|\cdot\|_{L^2(\Omega, H^{s+1}(D))}$  is the standard Bochner norm. Moreover, under the assumptions of Theorem 4.1, the following error bound on  $\mathbb{E}[\mathcal{Q}(u)]$  holds, with  $\mu = \min\{s, t, p\}$ :

$$\begin{aligned} & \|\mathbb{E}[\mathcal{Q}(u)] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)} \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left\| \left( \frac{a_{max}}{a_{min}} \right)^2 \right\|_{L^2(\Omega)} \|a_{min}^{-1}\|_{L^2(\Omega)} \left( \frac{h^{2\mu}}{p^{2 \min\{s, t-1\}}} + M^{-1/2} \right) \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right). \end{aligned} \quad (43)$$

*Proof.* The proof of (42) follows the same steps as the proof of [5, Theorem 4.3]. We include the proof here for completeness. By the triangle inequality, we get

$$\begin{aligned} & \|\mathbb{E}[u] - E_M[\Pi_{h,p}^{\nabla} u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))} \\ & \leq \underbrace{\|\mathbb{E}[u] - \mathbb{E}[\Pi_{h,p}^{\nabla} u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))}}_{(I)} + \underbrace{\|\mathbb{E}[\Pi_{h,p}^{\nabla} u_{h,p}] - E_M[\Pi_{h,p}^{\nabla} u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))}}_{(II)}. \end{aligned}$$

We estimate (I) by using the approximation and stability properties (14) and (17) of  $\Pi_{h,p}^{\nabla}$ , as well as the error estimate (25) of the VE method:

$$\begin{aligned} \|\mathbb{E}[u] - \mathbb{E}[\Pi_{h,p}^{\nabla} u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))} & = \|\mathbb{E}[u - \Pi_{h,p}^{\nabla} u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))} \\ & \leq \mathbb{E}[|u - \Pi_{h,p}^{\nabla} u_{h,p}|_{H^1(\mathcal{T}_h)}] \\ & \leq \mathbb{E}[|u - \Pi_{h,p}^{\nabla} u|_{H^1(\mathcal{T}_h)}] + \mathbb{E}[|\Pi_{h,p}^{\nabla}(u - u_{h,p})|_{H^1(\mathcal{T}_h)}] \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left\| \frac{a_{max}}{a_{min}} \right\|_{L^2(\Omega)} \frac{h^{\mu_s}}{p^s} \|u\|_{L^2(\Omega, H^{s+1}(D))}. \end{aligned}$$

For (II), we use the MC error estimate (38) and the stability estimates (17) and (24) obtaining

$$\begin{aligned} (II) & \leq M^{-1/2} \|\Pi_{h,p}^{\nabla} u_{h,p}\|_{L^2(\Omega, H^1(\mathcal{T}_h))} \lesssim M^{-1/2} \|u_{h,p}\|_{L^2(\Omega, H^1(\mathcal{T}_h))} \\ & \lesssim s_*(p)^{-1} \|a_{min}^{-1}\|_{L^2(\Omega)} M^{-1/2} \|f\|_{L^2(D)}, \end{aligned}$$

from which (42) follows, taking into account that  $s_*(p)^{-1} \leq \frac{s^*(p)}{s_*(p)}$ .

The error bound (43) follows similarly, by adding and subtracting the term  $\mathbb{E}[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]$  into  $\|\mathbb{E}[\mathcal{Q}(u)] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u)]\|_{L^2(\Omega)}$ :

$$\begin{aligned} & \|\mathbb{E}[\mathcal{Q}(u)] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)} \\ & \leq \underbrace{\|\mathbb{E}[\mathcal{Q}(u)] - \mathbb{E}[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)}}_{(I)} + \underbrace{\|\mathbb{E}[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)}}_{(II)}. \end{aligned}$$

Thanks to the linearity of the operator  $\mathbb{E}[\cdot]$ , and using  $\mathcal{Q}(\Pi_{h,p}^0 u_{h,p}) = \mathcal{Q}_h(u_{h,p})$  (see (27)), there holds:

$$\begin{aligned} (I) & = \|\mathbb{E}[\mathcal{Q}(u) - \mathcal{Q}_h(u_{h,p})]\|_{L^2(\Omega)} = \mathbb{E}[\|\mathcal{Q}(u) - \mathcal{Q}_h(u_{h,p})\|] \\ & \lesssim \frac{s^*(p)}{s_*(p)} \mathbb{E} \left[ \left( \frac{a_{max}}{a_{min}} \right)^2 a_{min}^{-1} \right] \frac{h^{2\mu}}{p^{2 \min\{s,t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right), \end{aligned}$$

where, in the last step, we applied (41). From (38), the definition of  $\mathcal{Q}$ , the Cauchy-Schwarz inequality, and the  $L^2(D)$  stability of  $\Pi_{h,p}^0$ , we have

$$(II) \lesssim M^{-1/2} \|\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})\|_{L^2(\Omega)} \leq M^{-1/2} \|q\|_{L^2(D)} \|u_{h,p}\|_{L^2(\Omega, L^2(D))}.$$

Using the Poincaré inequality and the stability property in (24), we obtain

$$(II) \lesssim s_*(p)^{-1} \|a_{min}^{-1}\|_{L^2(\Omega)} M^{-1/2} \|q\|_{L^2(D)} \|f\|_{L^2(D)}.$$

The above estimates of (I) and (II) give (43), taking onto account that

$$\begin{aligned} & \max \left\{ \frac{s^*(p)}{s_*(p)} \mathbb{E} \left[ \left( \frac{a_{max}}{a_{min}} \right)^2 a_{min}^{-1} \right], s_*(p)^{-1} \|a_{min}^{-1}\|_{L^2(\Omega)} \right\} \\ & \leq \frac{s^*(p)}{s_*(p)} \left\| \left( \frac{a_{max}}{a_{min}} \right)^2 \right\|_{L^2(\Omega)} \|a_{min}^{-1}\|_{L^2(\Omega)}. \end{aligned}$$

□

**Remark 5.2.** *The regularity assumption in Proposition 5.1 is satisfied provided that the diffusion coefficient  $a$  is sufficiently smooth. For instance, if  $a(\omega, \cdot) \in W^{1,\infty}(D)$  satisfies  $\|a(\omega, \cdot)\|_{W^{1,\infty}(D)} \leq C$  uniformly over  $\omega \in \Omega$ , the elliptic regularity theory ensures that  $\|u(\omega, \cdot)\|_{H^2(D)}$  is bounded uniformly over  $\omega \in \Omega$ , and thus fulfills  $u \in L^2(\Omega, H^2(D))$ . This holds, for example, if  $a(\omega, x)$  is given by the following Karhunen–Loève-type expansion:*

$$a(\omega, x) = \bar{a}(x) + \sum_{j=1}^J \beta_j \varphi_j(x) Y_j(\omega),$$

where  $\bar{a} \in W^{1,\infty}(D)$  is uniformly positive,  $\{Y_j\}_{j=1}^J$  is a set of independent uniformly distributed random variables  $Y_j \sim \mathcal{U}([-1, 1])$ ,  $\{\varphi_j\}_{j=1}^J \subset W^{1,\infty}(D)$ , and  $\{\beta_j\}_{j=1}^J$  is a set of non-negative coefficients such that  $\beta_j \|\varphi_j\|_{W^{1,\infty}(D)} \leq C j^{-n}$  for some  $n > 1$  and constant  $C > 0$  independent of  $j$ .

For diffusion coefficients with  $\omega$ -dependent lower and upper bounds  $a_{min}(\omega)$ ,  $a_{max}(\omega)$ , proving that  $u \in L^2(\Omega, H^2(D))$  requires a more delicate analysis, since  $\|u(\omega, \cdot)\|_{H^2(D)}$  has an  $\omega$ -dependent upper bound. In [21, Theorem 2.1], it is proved that, if  $a \in L^p(\Omega, C^t(\bar{D}))$  and  $f \in H^{t-1}(D)$  for some  $0 < t \leq 1$  and any  $p < \infty$ , then  $u \in L^p(\Omega, H^2(D))$ . For instance, this applies when the diffusion coefficient in problem (1) is lognormal, namely,  $a(\omega, x) = \exp(Y(\omega, x))$ , with  $Y$  being a Gaussian random field whose mean is Hölder continuous and whose covariance function is Lipschitz continuous. ■

**Remark 5.3.** *The choice  $M = \mathcal{O}(p^{2s} h^{-2\mu_s})$  balances the MC error ( $\mathcal{O}(M^{-1/2})$ ) with the VE discretization error ( $\mathcal{O}(h^{\mu_s} p^{-s})$ ) in (42), resulting in the overall error of  $\mathcal{O}(h^{\mu_s} p^{-s} \frac{s^*(p)}{s_*(p)})$ . This is the optimal choice of  $M$ , given the smoothness of  $u$ , the mesh size  $h$  and the order  $p$ , to achieve a prescribed tolerance. Analogously, the choice  $M = \mathcal{O}(p^{4 \min\{s,t-1\}} h^{-4\mu})$  balances the MC error ( $\mathcal{O}(M^{-1/2})$ ) with the VE discretization error ( $\mathcal{O}(h^{2\mu} p^{-2 \min\{s,t-1\}} \frac{s^*(p)}{s_*(p)})$ ) in (43). ■*

**Remark 5.4.** *One could alternatively define a MC-VE estimator for  $\mathbb{E}[u]$  employing the  $L^2$  projector:*

$$E_M[\Pi_{h,p}^0 u_{h,p}] = \frac{1}{M} \sum_{i=1}^M \Pi_{h,p}^0 u_{h,p}^{(i)} \in L^2(\Omega, L^2(D)), \quad (44)$$

for which the following bound can be proved along the same steps as in the proof of Proposition 5.1:

$$\|\mathbb{E}[u] - E_M[\Pi_{h,p}^0 u_{h,p}]\|_{L^2(\Omega, L^2(D))} \lesssim \left( \frac{h^{\mu_s+1}}{p^{s+1}} + M^{-1/2} \right). \quad (45)$$

This estimator provides no control in the  $L^2(\Omega, H^1(\mathcal{T}_h))$  norm. The use of definition (44) is particularly helpful in applications where multiple functionals  $\mathcal{Q}$  of the solution of the form (4) are of interest. Indeed, noting that  $E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})] = \mathcal{Q}(E_M[\Pi_{h,p}^0 u_{h,p}])$ , one can compute the MC-VE estimator  $E_M[\Pi_{h,p}^0 u_{h,p}]$  for  $\mathbb{E}[u]$ , and then the MC-VE estimator for  $\mathbb{E}[\mathcal{Q}(u)]$  can be obtained by simply applying the functional  $\mathcal{Q}$  to  $E_M[\Pi_{h,p}^0 u_{h,p}]$ , for all occurrences of  $\mathcal{Q}$ , allowing for substantial computational savings.

For the error analysis, we observe that the following holds

$$\begin{aligned} \|\mathbb{E}[\mathcal{Q}(u)] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)} &= \|\mathcal{Q}(\mathbb{E}[u] - E_M[\Pi_{h,p}^0 u_{h,p}])\|_{L^2(\Omega)} \\ &\leq \|q\|_{L^2(D)} \|\mathbb{E}[u] - E_M[\Pi_{h,p}^0 u_{h,p}]\|_{L^2(\Omega, L^2(D))}, \end{aligned}$$

where the equality follows from Fubini's theorem and the linearity of  $\mathcal{Q}$ , and the inequality follows from the Cauchy-Schwarz inequality. Hence, by using (45), we could directly derive an upper bound of the error  $\|\mathbb{E}[\mathcal{Q}(u)] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)}$ . This bound, however, would not be sharp; cf. (43). ■

To alleviate the computational burden of the MC-VE method without compromising accuracy, we study the multi-level  $h$ -version of the MC-VE method (MLMC-VE) in the following section. This approach considers multiple spatial resolution levels: the estimator is first computed on a coarse mesh using the MC-VE method, and subsequent correction terms are added to improve accuracy.

## 6 The MLMC-VE method

The MLMC-VE method we introduce in this section relies on a sequence of nested polygonal meshes  $\{\mathcal{T}_\ell\}_{\ell \geq 0}$  discretizing the domain  $D$ :

$$\mathcal{T}_0 \subseteq \mathcal{T}_1 \subseteq \dots \subseteq \mathcal{T}_\ell \subseteq \dots \subseteq \mathcal{T}_L \subseteq \dots, \quad (46)$$

where each  $\mathcal{T}_\ell$  is a refinement of  $\mathcal{T}_{\ell-1}$ , and  $\mathcal{T}_\ell$  has meshsize  $h_\ell$ . Let  $V_\ell$  denote the VE space of degree  $p$  (fixed and independent of  $\ell$ ) on the mesh  $\mathcal{T}_\ell$ . The sequence of meshes (46) induces the sequence  $\{V_\ell\}_{\ell \geq 0}$  of conforming VE spaces, which are not nested, although their polynomial subspaces are. This structure guarantees improved approximation properties as  $\ell \rightarrow \infty$ , almost surely in  $\Omega$ :

$$\begin{aligned} \|u(\omega, \cdot) - u_{h_{\ell+1}, p}(\omega, \cdot)\|_V &\leq \|u(\omega, \cdot) - u_{h_\ell, p}(\omega, \cdot)\|_V, \\ \|u(\omega, \cdot) - u_{h_\ell, p}(\omega, \cdot)\|_V &\xrightarrow{\ell \rightarrow \infty} 0, \end{aligned}$$

where  $u(\omega, \cdot)$  denotes the solution of (2), and  $u_{h_{\ell+1}, p}(\omega, \cdot) \in V_\ell$  denotes its VE approximation, i.e., the solution of (23). Since  $p$  is fixed, to simplify the notation we set  $\Pi_\ell^\nabla = \Pi_{h_\ell, p}^\nabla$  and  $u_{h_\ell, p} = u_\ell$ . Given the finite sequence  $\{\mathcal{T}_\ell\}_{\ell=0}^L$  of the first  $L+1$  nested meshes from (46), and the corresponding sequence  $\{V_\ell\}_{\ell=0}^L$  of VE spaces, we define the MLMC-VE estimator for  $\mathbb{E}[u]$  as

$$E^L[u] = \sum_{\ell=1}^L E_{M_\ell}[w_\ell] = \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} w_\ell^{(i)} \in L^2(\Omega, H^1(\mathcal{T}_L)), \quad (47)$$

where

$$w_\ell = \Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}, \quad w_\ell^{(i)} = \Pi_\ell^\nabla u_\ell^{(i)} - \Pi_{\ell-1}^\nabla u_{\ell-1}^{(i)},$$

and  $u_0^{(i)} = 0$ . Note that the MLMC-VE estimator (47) consists of the sum over the  $L$  levels of the MC-VE estimators of the increments  $w_\ell$ , using a level-dependent number  $M_\ell$  of samples.

**Remark 6.1.** Since  $u_\ell^{(i)}$  and  $u_{\ell-1}^{(i)}$  are VE functions belonging to two different, non-nested VE spaces  $V_\ell$  and  $V_{\ell-1}$ , respectively, an MLMC-VE estimator of the type

$$E^L[u] = \sum_{\ell=1}^L E_{M_\ell}[u_\ell - u_{\ell-1}],$$

would not be computable. This is why, in our definition (47), we insert the projection operators  $\Pi_\ell^\nabla$  and  $\Pi_{\ell-1}^\nabla$ , which makes the estimator computable. ■

Similarly to (47), we define the MLMC-VE estimator for  $\mathbb{E}[\mathcal{Q}(u)]$  as

$$E^L[\mathcal{Q}(u)] = \sum_{\ell=1}^L E_{M_\ell}[\mathcal{Q}(\nu_\ell)] = \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \mathcal{Q}(\nu_\ell^{(i)}) \in L^2(\Omega), \quad (48)$$

where

$$\nu_\ell = \Pi_\ell^0 u_\ell - \Pi_{\ell-1}^0 u_{\ell-1}, \quad \nu_\ell^{(i)} = \Pi_\ell^0 u_\ell^{(i)} - \Pi_{\ell-1}^0 u_{\ell-1}^{(i)}.$$

In the next theorem, we establish error estimates for the MLMC-VE estimators (47) and (48).

**Theorem 6.1** (Error estimate for the MLMC-VE estimators). *Assume that  $u \in L^2(\Omega, H^{s+1}(D))$ ,  $s \geq 1$ . Then, we have, with  $\mu_s = \min\{s, p\}$*

$$\begin{aligned} & \|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left\| \frac{a_{max}}{a_{min}} \right\|_{L^2(\Omega)} p^{-s} \left( h_L^{\mu_s} + M_1^{-1/2} h_1^{\mu_s} + \sum_{\ell=2}^L M_\ell^{-1/2} (h_\ell^{\mu_s} + 2h_{\ell-1}^{\mu_s}) \right) \|u\|_{L^2(\Omega, H^{s+1}(D))}. \end{aligned} \quad (49)$$

Moreover, under the assumptions of Theorem 4.1, the following error bound holds, with  $\mu = \min\{s, t, p\}$ :

$$\begin{aligned} & \|\mathbb{E}[\mathcal{Q}(u)] - E^L[\mathcal{Q}(u)]\|_{L^2(\Omega)} \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left\| \left( \frac{a_{max}}{a_{min}} \right)^2 a_{min}^{-1} \right\|_{L^2(\Omega)} p^{-2 \min\{s, t-1\}} \left( h_L^{2\mu} + M_1^{-1/2} h_1^{2\mu} + \sum_{\ell=2}^L M_\ell^{-1/2} (h_\ell^{2\mu} + h_{\ell-1}^{2\mu}) \right) \\ & \quad \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right). \end{aligned} \quad (50)$$

*Proof.* We start proving (49). By the triangular inequality, we get

$$\begin{aligned} & \|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \leq \underbrace{\|\mathbb{E}[u] - \mathbb{E}[\Pi_L^\nabla u_L]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}}_{(I)} + \underbrace{\|\mathbb{E}[\Pi_L^\nabla u_L] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}}_{(II)}. \end{aligned}$$

For (I), proceeding as in the proof of Proposition 5.1, we obtain

$$(I) \lesssim \frac{s^*(p)}{s_*(p)} \left\| \frac{a_{max}}{a_{min}} \right\|_{L^2(\Omega)} \frac{h_L^{\mu_s}}{p^s} \|u\|_{L^2(\Omega, H^{s+1}(D))}.$$

We focus now on the second term (II). Expressing  $\Pi_L^\nabla u_L$  as the telescopic sum

$$\Pi_L^\nabla u_L = \sum_{\ell=1}^L (\Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}) = \sum_{\ell=1}^L w_\ell,$$

the linearity of the expectation, together with (47) and (38), yields

$$\begin{aligned} (II) & = \left\| \sum_{\ell=1}^L (\mathbb{E}[w_\ell] - E_{M_\ell}[w_\ell]) \right\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \leq \sum_{\ell=1}^L \|\mathbb{E}[w_\ell] - E_{M_\ell}[w_\ell]\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \leq \sum_{\ell=1}^L M_\ell^{-1/2} \|w_\ell\|_{L^2(\Omega, H^1(\mathcal{T}_L))}, \end{aligned}$$

where the last inequality follows from (38). We estimate  $\|w_\ell\|_{L^2(\Omega, H^1(\mathcal{T}_L))}$  as follows:

$$\begin{aligned} \|w_\ell\|_{L^2(\Omega, H^1(\mathcal{T}_L))} & = \|\Pi_\ell^\nabla (u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1})\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \lesssim \|u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \lesssim \|u_\ell - u\|_{L^2(\Omega, H^1(\mathcal{T}_L))} + \|u - \Pi_{\ell-1}^\nabla u\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \quad + \|\Pi_{\ell-1}^\nabla (u - u_{\ell-1})\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left\| \frac{a_{max}}{a_{min}} \right\|_{L^2(\Omega)} \frac{(h_\ell^{\mu_s} + 2h_{\ell-1}^{\mu_s})}{p^s} \|u\|_{L^2(\Omega, H^{s+1}(D))}, \end{aligned}$$

where we used the identity  $\Pi_\ell^\nabla(\Pi_{\ell-1}^\nabla u_{\ell-1}) = \Pi_{\ell-1}^\nabla u_{\ell-1}$ , which is a consequence of the fact that the meshes are nested, cf. (46), the triangle inequality, the stability and approximation properties of the projection operators, as well as the error estimate (25). Then, we obtain

$$(II) \lesssim \frac{s^*(p)}{s_*(p)} \left\| \frac{a_{max}}{a_{min}} \right\|_{L^2(\Omega)} \sum_{\ell=1}^L (M_\ell)^{-1/2} \frac{(h_\ell^{\mu_s} + 2h_{\ell-1}^{\mu_s})}{p^s} \|u\|_{L^2(\Omega, H^{s+1}(D))},$$

which concludes the proof of (49).

The proof of (50) works similarly. First, we sum and subtract  $\mathbb{E}[\mathcal{Q}(u)]$  and use the triangular inequality:

$$\begin{aligned} & \|\mathbb{E}[\mathcal{Q}(u)] - E^L[\mathcal{Q}(u)]\|_{L^2(\Omega, V)} \\ & \leq \underbrace{\|\mathbb{E}[\mathcal{Q}(u)] - \mathbb{E}[\mathcal{Q}(\Pi_L^0 u_L)]\|_{L^2(\Omega)}}_{(I)} + \underbrace{\|\mathbb{E}[\mathcal{Q}(\Pi_L^0 u_L)] - E^L[\mathcal{Q}(u)]\|_{L^2(\Omega)}}_{(II)}. \end{aligned}$$

The term (I) is the same as in the proof of Proposition 5.1, hence there holds

$$(I) \lesssim \frac{s^*(p)}{s_*(p)} \mathbb{E} \left[ \left( \frac{a_{max}}{a_{min}} \right)^2 a_{min}^{-1} \right] \frac{h_L^{2\mu}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right).$$

To estimate (II), we express  $\Pi_L^0 u_L$  as telescopic sum

$$\Pi_L^0 u_L = \sum_{\ell=1}^L (\Pi_\ell^0 u_\ell - \Pi_{\ell-1}^0 u_{\ell-1}) = \sum_{\ell=1}^L \nu_\ell,$$

and we use the linearity of the expectation, the linearity of  $\mathcal{Q}$ , together with (48) and (38):

$$(II) \leq \sum_{\ell=1}^L M_\ell^{-1/2} \|\mathcal{Q}(\nu_\ell)\|_{L^2(\Omega)}.$$

Finally, we estimate  $\|\mathcal{Q}(\nu_\ell)\|_{L^2(\Omega)}$  as follows:

$$\begin{aligned} \|\mathcal{Q}(\nu_\ell)\|_{L^2(\Omega)} & = \|\mathcal{Q}(\Pi_\ell^0 u_\ell - \Pi_{\ell-1}^0 u_{\ell-1})\|_{L^2(\Omega)} \\ & \leq \|\mathcal{Q}(\Pi_\ell^0 u_\ell - u)\|_{L^2(\Omega)} + \|\mathcal{Q}(u - \Pi_{\ell-1}^0 u_{\ell-1})\|_{L^2(\Omega)} \\ & \lesssim \frac{s^*(p)}{s_*(p)} \left\| \left( \frac{a_{max}}{a_{min}} \right)^2 a_{min}^{-1} \right\|_{L^2(\Omega)} \frac{h_\ell^{2\mu} + h_{\ell-1}^{2\mu}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right), \end{aligned}$$

where we used (41). Then, we obtain

$$(II) \lesssim \frac{s^*(p)}{s_*(p)} \left\| \left( \frac{a_{max}}{a_{min}} \right)^2 a_{min}^{-1} \right\|_{L^2(\Omega)} \sum_{\ell=1}^L (M_\ell)^{-1/2} \frac{h_\ell^{2\mu} + h_{\ell-1}^{2\mu}}{p^{2 \min\{s, t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right),$$

which concludes the proof of (50).  $\square$

**Remark 6.2.** Employing the MC-VE estimator defined in (44) in the construction of  $E^L[u]$  (cf. (47)), one could easily adapt the above proof and obtain an estimate for the error  $\|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, L^2(D))}$ .  $\blacksquare$

The previous theorem gives the error estimate for the MLMC-VE method for any distribution of number of samples per level  $\{M_\ell\}_{\ell=1}^L$ . The following result, being a multilevel version of Remark 5.3, specifies how to choose the number of samples  $M_\ell$  at each mesh level  $\ell$  to achieve an overall error of order  $\mathcal{O}(h_L^{\mu_s} p^{-s} \frac{s^*(p)}{s_*(p)})$  for the MLMC-VE estimator.

**Theorem 6.2.** Consider the same assumptions as in Theorem 6.1. Moreover, assume, for simplicity, that  $h_{\ell-1} = 2h_\ell$  for all  $\ell = 1, \dots, L$ . Then, the MLMC-VE estimator  $E^L[u]$  with

$$M_\ell \sim \left( h_\ell^{\mu_s} h_L^{-\mu_s} \ell^{(1+\varepsilon)} \right)^2 \quad (51)$$

samples (with  $\varepsilon > 0$  arbitrarily small) on the mesh level  $\ell$  admits the error bound

$$\|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \lesssim \frac{s^*(p)}{s_*(p)} \frac{h_L^{\mu_s}}{p^s} \|u\|_{L^2(\Omega, H^{s+1}(D))}. \quad (52)$$

Similarly, for the MLMC-VE estimator  $E^L[\mathcal{Q}(u)]$ , with

$$M_\ell = \left( h_\ell^{2\mu} h_L^{-2\mu} \ell^{(1+\varepsilon)} \right)^2$$

samples on the mesh level  $\ell$ , the following bound holds:

$$\|\mathbb{E}[\mathcal{Q}(u)] - E^L[\mathcal{Q}(u)]\|_{L^2(\Omega)} \lesssim \frac{s^*(p)}{s_*(p)} \frac{h_L^{2\mu}}{p^{2\min\{s,t-1\}}} \left( \|q\|_{H^{t-1}(D)}^2 + \|f\|_{H^{s-1}(D)}^2 \right).$$

*Proof.* With  $h_{\ell-1} = 2h_\ell$  and  $M_\ell$  as in (51), we get

$$\begin{aligned} M_1^{-1/2} h_1^{\mu_s} + \sum_{\ell=2}^L M_\ell^{-1/2} (h_\ell^{\mu_s} + 2h_{\ell-1}^{\mu_s}) &= M_1^{-1/2} h_1^{\mu_s} + \sum_{\ell=2}^L M_\ell^{-1/2} (h_\ell^{\mu_s} + 2 \cdot 2^{\mu_s} h_\ell^{\mu_s}) \\ &= M_1^{-1/2} h_1^{\mu_s} + (1 + 2^{\mu_s+1}) \sum_{\ell=2}^L M_\ell^{-1/2} h_\ell^{\mu_s} \sim h_1^{-\mu_s} h_L^{\mu_s} h_1^{\mu_s} + (1 + 2^{\mu_s+1}) \sum_{\ell=2}^L \frac{h_L^{\mu_s}}{h_\ell^{\mu_s}} \ell^{-(1+\varepsilon)} h_\ell^{\mu_s} \\ &< (1 + 2^{\mu_s+1}) h_L^{\mu_s} \sum_{\ell=1}^L \ell^{-(1+\varepsilon)} = C(\varepsilon) (1 + 2^{\mu_s+1}) h_L^{\mu_s}, \end{aligned}$$

where  $C(\varepsilon)$  is a positive constant depending on  $\varepsilon$  (more precisely,  $C(\varepsilon) = \zeta(1 + \varepsilon)$ , with  $\zeta$  the Riemann zeta function). Inserting this into (49) gives (52). The result for the QoI can be proved with a similar argument.  $\square$

A practical choice for the sample sizes, which avoids requiring knowledge of the solution's regularity, is discussed in Section 7.2 below.

**Remark 6.3.** For the MLMC-VE estimator, we write

$$\begin{aligned} \|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}^2 &= \|\mathbb{E}[u - \Pi_L^\nabla u_L]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}^2 + \mathbb{V}[E^L[u]] \\ &= \|\mathbb{E}[u - \Pi_L^\nabla u_L]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}^2 + \sum_{\ell=1}^L \frac{\mathbb{V}[\Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}]}{M_\ell}, \end{aligned} \quad (53)$$

where  $\mathbb{V}[E^L[u]] := \|E^L[u] - \mathbb{E}[E^L[u]]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}^2$ . The first identity follows from  $\mathbb{E}[E^L[u]] = \mathbb{E}[\Pi_L^\nabla u_L]$ , which implies

$$\mathbb{V}[E^L[u]] = \|E^L[u] - \mathbb{E}[\Pi_L^\nabla u_L]\|_{L^2(\Omega, H^1(\mathcal{T}_L))}^2,$$

while the second identity follows from

$$\mathbb{V}[E_{M_\ell}[\Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}]] = \frac{1}{M_\ell} \mathbb{V}[\Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}].$$

The error split in (53) is known as bias–variance decomposition. The first term is the bias component of the error, which is due to the VE discretization error at the finest level  $L$ . The second term is the variance component of the error, which is due to the sampling error. Assume that one wants to define the estimator  $E^L[u]$  by selecting the finest level  $L$  and the number of samples  $M_\ell$  per level such that

$$\|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))} \leq \varepsilon,$$

for a prescribed tolerance  $\varepsilon > 0$ . To achieve this, the tolerance is typically distributed between the bias and the variance:

$$\varepsilon^2 = \varepsilon_{\text{bias}}^2 + \varepsilon_{\text{var}}^2,$$

where, for  $\theta \in (0, 1)$ ,  $\varepsilon_{\text{bias}}^2 = \theta\varepsilon^2$  controls the space discretization error and  $\varepsilon_{\text{var}}^2 = (1 - \theta)\varepsilon^2$  controls the sampling error. An extensive discussion on the optimal choice of the splitting parameter can be found in [19]. The most common choice is  $\theta = 1/2$ , yielding an equal splitting of the total tolerance between the bias and variance components. Following the bias–variance decomposition, the finest level  $L$  is chosen so that the discretization error is below  $\varepsilon_{\text{bias}}$ , and the number of samples  $M_\ell$  at each level  $\ell$  is selected so as to minimize the variance contribution to the total error for a given computational cost. This leads to

$$M_\ell \propto \sqrt{\frac{\mathbb{V}[\Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}]}{C_\ell}},$$

where  $C_\ell$  denotes the cost of computing a single realization of  $\Pi_\ell^\nabla u_\ell - \Pi_{\ell-1}^\nabla u_{\ell-1}$  (see, e.g., [14]).  $\blacksquare$

## 7 Numerical results

In this section, we first verify the convergence estimate of the VE error for a linear QoI in the deterministic case, as established in Section 4. We then proceed to validate

### 7.1 Verification of the estimates of the VE error for the QoI

Let us consider problem (5) on  $D = (0, 1) \times (0, 1) \subset \mathbb{R}^2$  with  $\alpha = 1$  and  $f(x_1, x_2) = 32\pi^2 u_{ex}(x_1, x_2)$ , whose analytical solution is  $u_{ex}(x_1, x_2) = \sin(4\pi x_1) \sin(4\pi x_2)$ . We use the sequence of non-nested meshes represented in Figure 1. Taking  $q(x) = 1$  in the definition of the quantity of interest (4), we get  $\mathcal{Q}(u_{ex}) = 0$ . We compute the approximation  $\mathcal{Q}_h(u_h)$  using the VE method of order  $p = 1, 2, 3$ . The resulting errors represented in Figure 2 demonstrate that  $|\mathcal{Q}(u_{ex}) - \mathcal{Q}_h(u_h)| = \mathcal{O}(h^{2p})$ , as proved in Theorem 4.1.

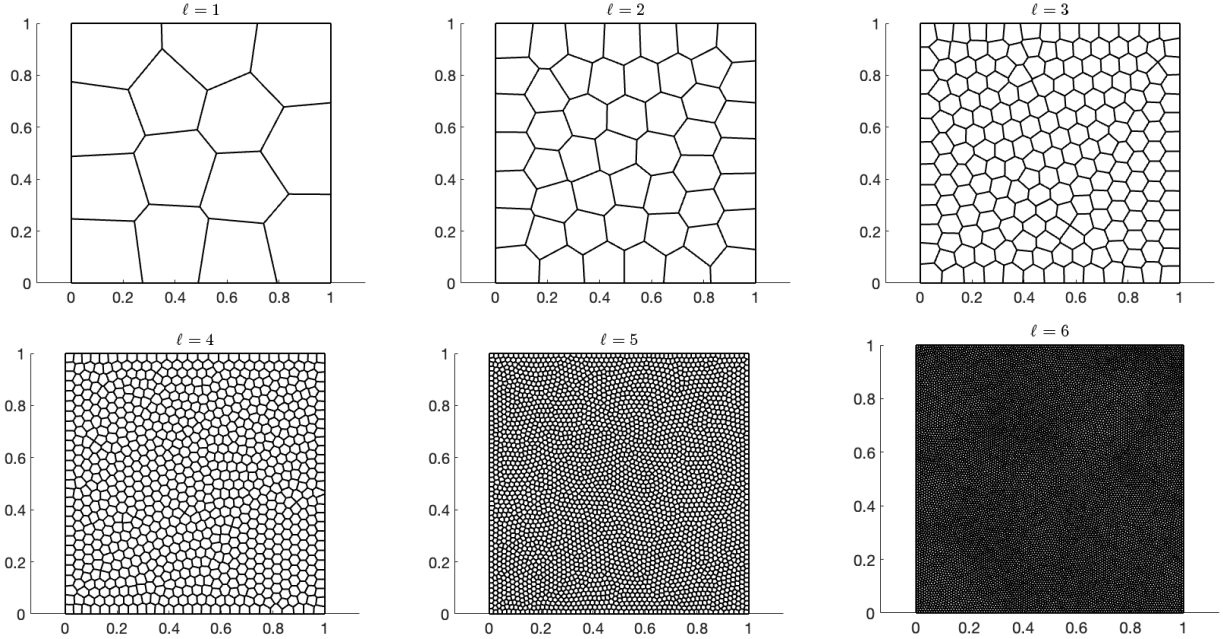


Figure 1: Sequence of non-nested Voronoi meshes used in the numerical tests of Section 7.1.

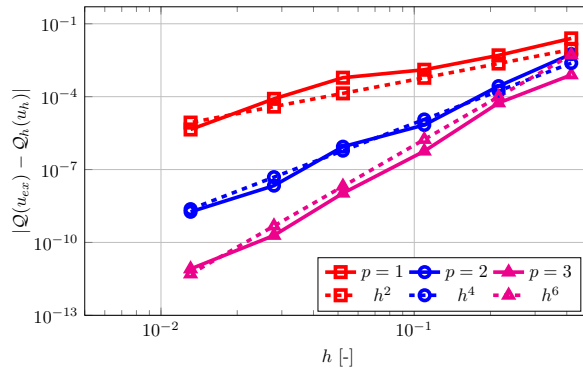


Figure 2: VE error for the QoI,  $|\mathcal{Q}(u_{ex}) - \mathcal{Q}_h(u_h)|$ , for  $p = 1, 2, 3$ .

### 7.2 Practical selection of MC and MLMC sample sizes

We discuss the practical selection of MC and MLMC sample sizes, taking into account the comparison between the MC-VE and MLMC-VE methods. Since the smoothness of the solution  $u$  may not be known a priori, we adopt the following choice of the number  $M$  of MC samples:

$$M = \begin{cases} p^{2p} h^{-2p}, & \text{for } E_M[u], \\ p^{4p} h^{-4p}, & \text{for } E_M[\mathcal{Q}(u)]. \end{cases} \quad (54)$$

Under the assumptions in the previous sections, this choice ensures the error bounds

$$\begin{aligned}\|\mathbb{E}[u] - E_M[\Pi_{h,p}^\nabla u_{h,p}]\|_{L^2(\Omega, H^1(\mathcal{T}_h))} &= \mathcal{O}(h^{\mu_s}), \\ \|\mathbb{E}[\mathcal{Q}(u)] - E_M[\mathcal{Q}(\Pi_{h,p}^0 u_{h,p})]\|_{L^2(\Omega)} &= \mathcal{O}(h^{2\mu}),\end{aligned}$$

where we recall the definitions  $\mu_s = \min\{s, p\}$  and  $\mu = \min\{s, t, p\}$ .

Similarly, the practical choice we adopt for the number of MLMC samples  $M_\ell$  at the level  $\ell$  is

$$M_\ell = \begin{cases} \left( h_\ell^p h_L^{-p} \ell^{(1+\varepsilon)} \right)^2, & \text{for } E_M[u], \\ \left( h_\ell^{2p} h_L^{-2p} \ell^{(1+\varepsilon)} \right)^2, & \text{for } E_M[\mathcal{Q}(u)]. \end{cases} \quad (55)$$

ensuring the error bounds

$$\begin{aligned}\|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, H^1(\mathcal{T}_L))} &= \mathcal{O}(h_L^{\mu_s}), \\ \|\mathbb{E}[\mathcal{Q}(u)] - E^L[\mathcal{Q}(u)]\|_{L^2(\Omega)} &= \mathcal{O}(h_L^{2\mu}).\end{aligned}$$

If the solution has high regularity, then the order of convergence  $\mu_s$  of the MC-VE and MLMC-VE errors on  $\mathbb{E}[u]$  is equal to  $p$ , and the order of convergence  $2\mu$  of the MC-VE and MLMC-VE errors on  $\mathbb{E}[\mathcal{Q}(u)]$  is equal to  $2p$ , provided that the function  $q$  defining the QoI is sufficiently smooth. However, if the solution is expected to have low regularity (and  $q$  has low regularity, in case of  $\mathbb{E}[\mathcal{Q}(u)]$ ), the use of high orders  $p$  is typically not beneficial. Therefore, we limit the following discussion to the case  $p \leq s$  (and  $p \leq t$ , in case of  $\mathbb{E}[\mathcal{Q}(u)]$ ). Given the practical choice (55), we estimate the computational complexity of the algorithm for the computation of the MLMC-VE estimator of  $\mathbb{E}[u]$  as follows.

**Theorem 7.1.** *Let  $N_\ell$  denote the number of degrees of freedom of the VE space of degree  $p$  on the mesh  $\mathcal{T}_\ell$ , and assume that  $N_\ell = \mathcal{O}(2^{2\ell} p^2)$  (this holds, for example, for Cartesian meshes). Moreover, assume that, at each level  $\ell$ , the VE equations for each sample  $u_\ell^{(i)}$  are solved employing a multigrid method, and that  $M_\ell$  is chosen according to (55). Then, the complexity  $\mathbb{C}_{MLMC}(L)$  for computing the MLMC-VE estimator  $E^L[u]$  ensuring an accuracy of order  $\mathcal{O}(h_L^p)$  satisfies*

$$\mathbb{C}_{MLMC}(L) \lesssim \begin{cases} N_L L^{3+2\varepsilon}, & \text{if } p = 1, \\ N_L^p p^{2(1-p)}, & \text{if } p \geq 2. \end{cases} \quad (56)$$

*Proof.* Choosing  $M_\ell$  as in (55), using  $h_\ell \sim 2^{-\ell}$  and assuming that the complexity of the multigrid algorithm at level  $\ell$  is proportional to  $N_\ell$ , then there holds

$$\begin{aligned}\mathbb{C}_{MLMC}(L) &\lesssim \sum_{\ell=1}^L M_\ell N_\ell \sim \sum_{\ell=1}^L \left( h_\ell^p h_L^{-p} \ell^{(1+\varepsilon)} \right)^2 N_\ell \sim \sum_{\ell=1}^L \left( 2^{p(L-\ell)} \ell^{(1+\varepsilon)} \right)^2 2^{2\ell} p^2 \\ &= p^2 2^{2L} \sum_{\ell=1}^L 2^{2p(L-\ell)} 2^{2(\ell-L)} \ell^{2(1+\varepsilon)} \sim N_L \sum_{\ell=1}^L 2^{(\ell-L)(2-2p)} \ell^{2(1+\varepsilon)}\end{aligned} \quad (57)$$

We now estimate the sum in (57).

- For  $p = 1$ , there holds

$$\sum_{\ell=1}^L 2^{(\ell-L)(2-2p)} \ell^{2(1+\varepsilon)} = \sum_{\ell=1}^L \ell^{2(1+\varepsilon)} = \sum_{\ell'=0}^{L-1} (L - \ell')^{2(1+\varepsilon)} = \mathcal{O}(L^{3+2\varepsilon}).$$

- For  $p \geq 2$ , then  $2 - 2p$  is negative and we call it  $-k$  with  $k > 0$ . Then, there holds

$$\begin{aligned}\sum_{\ell=1}^L 2^{(\ell-L)(2-2p)} \ell^{2(1+\varepsilon)} &= \sum_{\ell=1}^L 2^{k(L-\ell)} \ell^{2(1+\varepsilon)} = \sum_{\ell'=0}^{L-1} (2^k)^{\ell'} (L - \ell')^{2(1+\varepsilon)} \\ &= \mathcal{O}(2^{kL}) = \mathcal{O}(2^{2L(p-1)}) = \mathcal{O}((N_L p^{-2})^{(p-1)}).\end{aligned}$$

Hence, (56) follows.  $\square$

From Theorem 7.1, the computational advantage of the MLMC-VE method with respect to the MC-VE method on the finest mesh  $\mathcal{T}_L$  is evident. Indeed, choosing  $M$  as in (54) and using  $N_L = \mathcal{O}(2^{2L}p^2)$ , the complexity  $\mathbb{C}_{MC}(L)$  for computing the MC-VE estimator of  $\mathbb{E}[u]$  to ensure an accuracy of order  $\mathcal{O}(h_L^p)$  is estimated as

$$\mathbb{C}_{MC}(L) = MN_L = \mathcal{O}(p^{2p}h_L^{-2p})N_L = \mathcal{O}(N_L^{p+1})$$

Comparing (56) in Theorem 7.1 with the above result indicates that the MLMC-VE method achieves the same accuracy of MC-VE with a substantial reduction of the complexity in the considered case  $p \leq s$ .

For the QoI, following similar steps as above, we can show that the practical choice of the number of samples  $M_\ell$  in (55) for the computation of the MLMC-VE estimator of  $\mathbb{E}[\mathcal{Q}(u)]$  results into a computational complexity  $\mathcal{O}(N_L^{2p}p^{2(1-2p)})$ , independently of whether  $p = 1$  or  $p \geq 2$ , which is substantially smaller than the complexity of computing the MC-VE estimator  $E_M[\mathcal{Q}(u)]$  on the finest mesh  $\mathcal{T}_L$  with the practical choice of  $M$  as in (54), which is  $\mathcal{O}(N_L^{2p+1})$ .

### 7.3 Verification of the error estimates for the MLMC-VE method

Let us consider problem (1) defined on  $D = (0, 1) \times (0, 1) \subset \mathbb{R}^2$  with constant forcing term  $f(\omega, x_1, x_2) = 1$  and random diffusion coefficient  $a(\omega, x_1, x_2) = 5 + x_1 + x_2 + \left(\frac{8}{\pi^2}\right)^{2.5} Y(\omega) \sin\left(\frac{\pi}{4}(x_1 + 1)\right) \sin\left(\frac{\pi}{4}(x_2 + 1)\right)$ , where  $Y \sim \mathcal{U}([-1, 1])$  is a uniformly distributed random variable. This example is taken from [5, Section 6]. With the aim of approximating  $\mathbb{E}[u]$  by means of the MLMC-VE estimator, we consider a sequence of nested meshes made of square elements with side length  $1/2^\ell$ , ( $h_\ell = \frac{\sqrt{2}}{2^{\ell+1}}$ ), for  $\ell = 1, \dots, 6$ . In Figure 3, we show the number of samples  $M_\ell$  per mesh level  $\ell$  required to achieve  $\|\mathbb{E}[u] - E^L[u]\|_{L^2(\Omega, V_b)} = \mathcal{O}(h_L^p)$  for  $L = 6$ , as predicted by Theorem 7.1. We recall that  $E^L[u] = \sum_{\ell=1}^L E_{M_\ell}[w_\ell]$ , see (47). In particular, Figure 3 (left) displays in semilog scale the number of samples per level

$$\left\{ M_\ell = \left( h_\ell^p h_L^{-p} \ell^{(1+\varepsilon)} \right)^2, \ell = 1, \dots, 6 \right\}$$

needed to compute the MLMC estimator  $E^L[u]$  with  $L = 6$  and  $\varepsilon = 1e - 10$ , for different values of the order  $p = 1, 2, 3$ . In Figure 3 (right), we plot in semilog scale the total number of samples  $\sum_{\ell=1}^L M_\ell$  versus the maximum level  $L$  for  $L = 1, \dots, 6$ . As expected, the number of samples increases as the maximum level  $L$  increases, and the rate of increase grows with the order  $p$ .

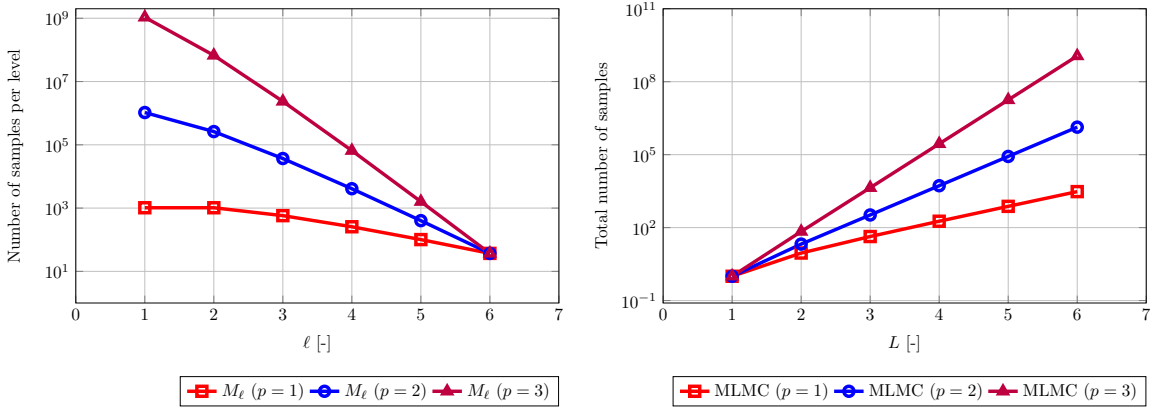


Figure 3: (Left) Number of samples per level  $\{M_\ell\}_{\ell=1}^6$  needed to compute the MLMC-VE estimator  $E^L[u] = \sum_{\ell=1}^L E_{M_\ell}[w_\ell]$  with  $L = 6$  according to (55) with  $\varepsilon = 1e - 10$ . (Right) Total number of samples  $\sum_{\ell=1}^L M_\ell$  needed to compute the MLMC estimator  $E^L[u]$  for increasing  $L = 2, \dots, 6$ . Both figures are in semilog scale.

For the orders  $p = 1, 2$ , we compute the MLMC-VE estimator  $E^L[u]$  for increasing values of  $L$  ( $L = 1, \dots, 5$  for  $p = 1$  and  $L = 1, \dots, 4$  for  $p = 2$ , respectively). Since the analytical solution  $\mathbb{E}[u]$  is unknown, for each  $p$ , we consider, as the reference solution, the MLMC-VE estimator with the maximum level  $L + 1$  (i.e., maximum level 6 for  $p = 1$  and maximum level 5 for  $p = 2$ ). In Figure 4 (left), we plot in loglog scale the  $H^1(D)$  norm of the errors of the MC-VE method and MLMC-VE method versus  $h_L$ . The theoretical estimates obtained in Sections 5 and 6 are confirmed (see Proposition 5.1 and Remark 5.3 for the MC-VE method, and Theorem 6.2 for the MLMC-VE method). Indeed, the plots clearly show that both the MC-VE and MLMC-VE errors converge with order  $\mathcal{O}(h_L^p)$ :

$$\|\mathbb{E}[u] - E_M[\Pi_{h,p}^\nabla u_h]\|_{H^1(D)} = \mathcal{O}(h_L^p), \quad \|\mathbb{E}[u] - E^L[u]\|_{H^1(D)} = \mathcal{O}(h_L^p).$$

Figure 4 (right) reports the MLMC-VE errors versus the maximum level  $L$  in semilog scale, for  $L = 1, \dots, 6$ . The results shows exponential decay in  $L$  with approximate slope of 0.8 for  $p = 1$  and 1.5 for  $p = 2$ . In the same setting, to assess the relative accuracy of MLMC-VE and MC-VE, we fix several levels of computational complexity and plot in loglog scale the corresponding errors for each method in Figure 5. We recall that, with the choices and the notation of Section 7.2, the computational complexity  $\mathcal{C}_{MLMC}(L_{MLMC})$  for MLMC with maximum level  $L_{MLMC}$  is, up to a constant,  $\sum_{\ell=1}^{L_{MLMC}} M_\ell N_\ell$ , and the computational complexity  $\mathcal{C}_{MC}(L_{MC})$  for MC at level  $L_{MC}$  is  $M N_{L_{MC}}$ . For a fixed value of  $\mathcal{C}_{MLMC}$ , the accuracy of the MLMC-VE method improves as  $p$  increases. In contrast, for a fixed  $\mathcal{C}_{MC}$ , the MC-VE method produces solutions with a level of accuracy independent of  $p$ . More importantly, at a given computational complexity, MLMC-VE consistently achieves higher accuracy, yielding errors roughly one order of magnitude smaller than those of MC-VE.

Finally, in Figure 6 (left), we plot in loglog scale the MLMC-VE error on  $\mathbb{E}[Q(u)]$  for the choice  $Q(u) = \int_D u \, dx$ . The number of samples  $M_\ell$  per level is fixed according to (55), and the predicted rate of convergence

$$|\mathbb{E}[Q(u)] - E^L[Q(u)]| = \mathcal{O}(h_L^{2p}).$$

is confirmed. The same errors are plotted in loglog scale versus the computational complexity in Figure 6 (right), where it is again evident that MLMC-VE consistently achieves higher accuracy than MC-VE.

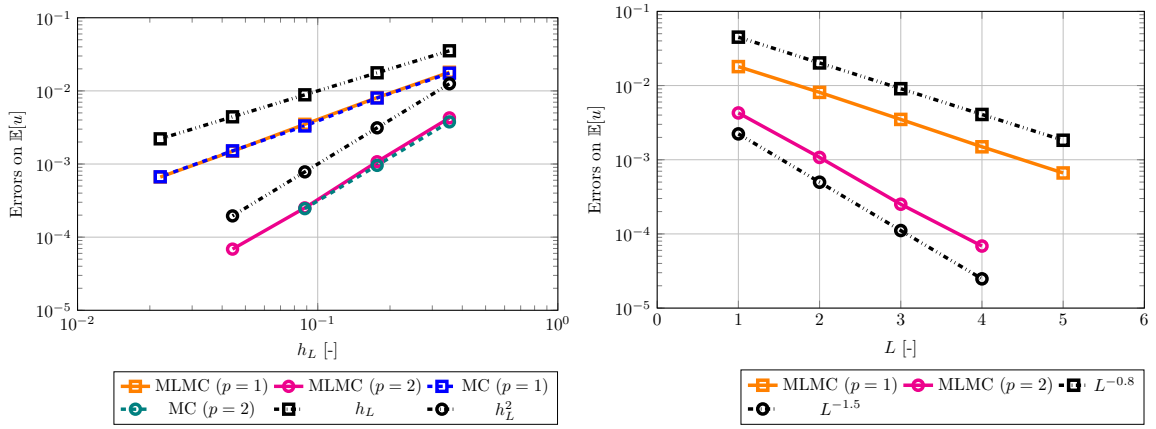


Figure 4: (Left) MC-VE and MLMC-VE errors on  $\mathbb{E}[u]$  in the  $H^1(D)$  norm, plotted versus  $h_L$  in loglog scale. (Right) MLMC-VE errors on  $\mathbb{E}[u]$  in the  $H^1(D)$  norm, plotted versus  $L$  in semilog scale.

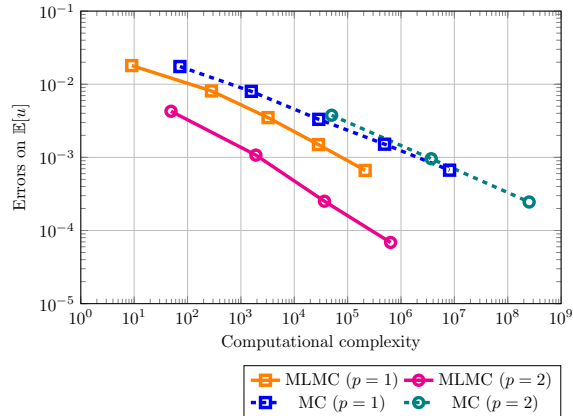


Figure 5: MLMC-VE and MC-VE errors on  $\mathbb{E}[u]$  in the  $H^1(D)$  norm versus the computational complexity in loglog scale.

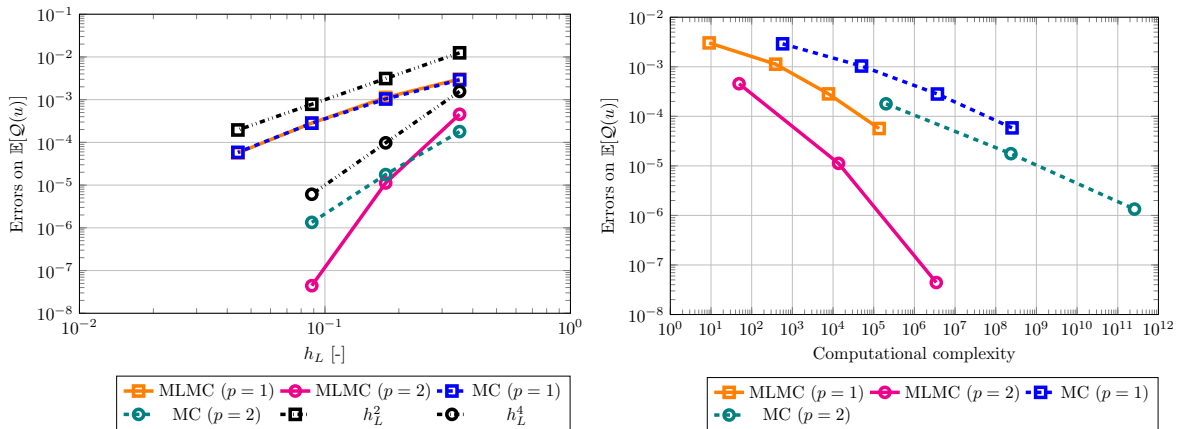


Figure 6: (Left) MLMC-VE errors on  $\mathbb{E}[Q(u)]$  plotted versus  $h_L$  in loglog scale. (Right) MLMC-VE errors on  $\mathbb{E}[Q(u)]$  versus the computational complexity in loglog scale.  $Q(u) = \int_D u \, dx$ .

## 7.4 Validation test

In this section, we present some numerical results to demonstrate the practical capabilities of the proposed scheme. We consider problem (1) defined over the rectangular domain  $D = (0, 4) \times (0, 1)$ , which is split into seven non-overlapping regions  $\{D_r\}_{r=1}^7$  (see Figure 7) modeling rock strata in the subsurface. We take the forcing term constant and equal to 1. The random diffusion coefficient is defined as  $a(\omega, x) = \sum_{r=1}^7 \chi_{D_r}(x) Y_r(\omega)$ , where  $\{Y_r\}_{r=1}^7$  are uniformly distributed random variables. In particular,

$$a(\omega, x)|_{D_r} = Y_r(\omega) \in \mathbb{R} \quad \text{for all } r = 1, \dots, 7,$$

so that the realizations of the random diffusion coefficient are piecewise constant over the regions, with each local coefficient following a uniform distribution. This example goes beyond the theoretical framework considered in the previous sections, as the solution of problem (1) with a discontinuous diffusion coefficient does not satisfy the regularity assumption  $u \in L^p(\Omega, H^2(D))$  (see Proposition 5.1 and Theorem 6.1). The MLMC-VE method



Figure 7: Subdivision of the rectangular domain into subregions, as considered in Section 7.4.

is defined on the sequence of four nested polygonal meshes  $\{\mathcal{T}_\ell\}_{\ell=0}^3$  represented in Figure 8. The sequence satisfies  $h_\ell \simeq \frac{h_{\ell-1}}{2}$  and is constructed using a mesh-agglomeration strategy. Specifically, starting from the fine (triangular) mesh  $T_L = \mathcal{T}_3$ , at each level  $\ell = L - 1, \dots, 0$ , we recursively group fine-level elements into coarse agglomerates so that each coarse cell is the union of a set of fine cells. For the numerical simulations, we use the algorithm proposed in [4, 3] and implemented in MAGNET, an open-source Python library for mesh agglomeration in two and three dimensions based on Graph Neural Networks (see [2]).

We first consider the case  $Y_r \sim \mathcal{U}([1, 10])$  for all  $r = 1, \dots, 7$ , modeling the setting where the diffusivity of all regions is comparable (ranging between 1 and 10), namely, all regions are composed of the same material. In Figure 9, we compare the MLMC estimator (top) with the MC estimator (middle) of  $\mathbb{E}[u]$ . The pointwise absolute difference is plotted in Figure 9 (bottom), showing agreement between the two approximations up to the third decimal place. The same quantities are then computed for  $Y_r \sim \mathcal{U}([1, 2])$  for all  $r \neq 7$  and  $Y_7 \sim \mathcal{U}([1, 100])$ . This configuration models a case in which the first six regions have comparable diffusivity (varying within the

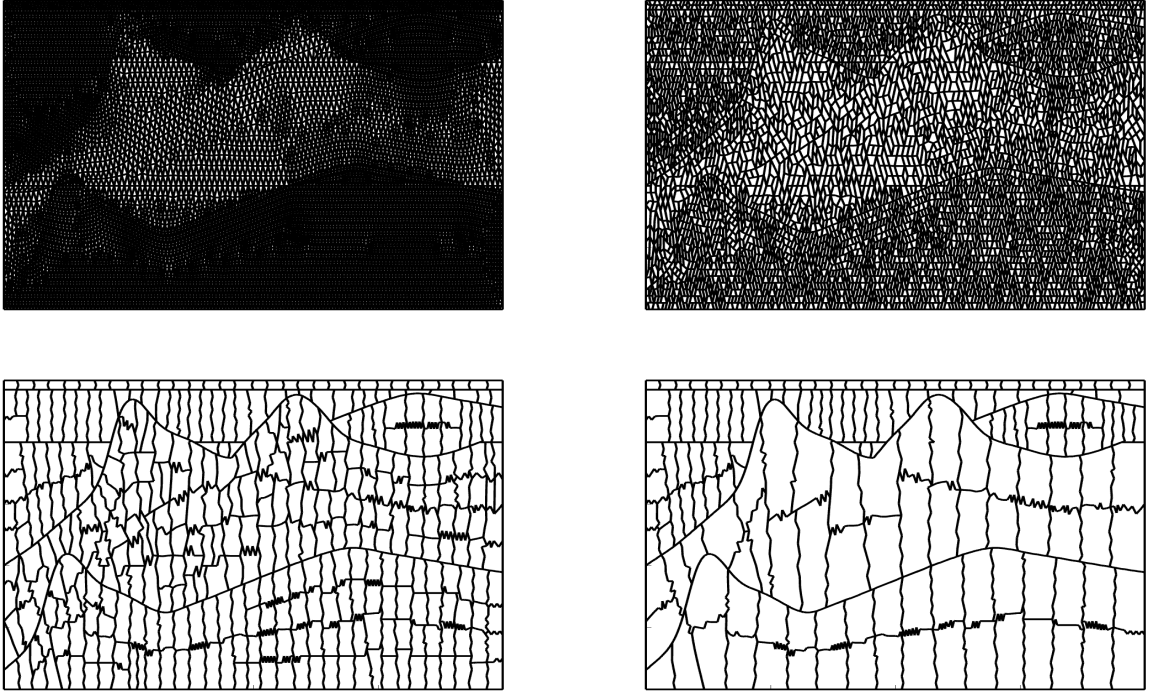


Figure 8: Sequence of nested meshes used in the numerical tests of Section 7.4, obtained via a mesh-agglomeration strategy from a fine triangular grid shown in the top-right panel.

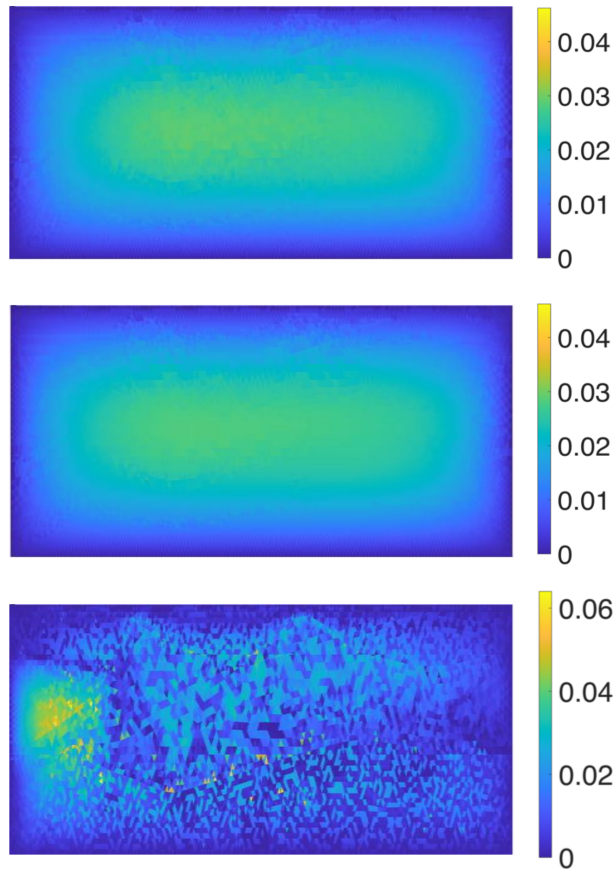


Figure 9: (Top): MLMC-VE approximation of  $\mathbb{E}[u]$  for  $a(\omega, x) = \sum_{r=1}^7 \chi_{D_r}(x) Y_r(\omega)$ , where  $Y_r \sim \mathcal{U}([1, 10])$  for all  $r = 1, \dots, 7$ . (Middle): MC approximation of  $\mathbb{E}[u]$ . (Bottom): Absolute value of the difference of the two approximations divided by the maximum norm of the MLMC-VE approximation.

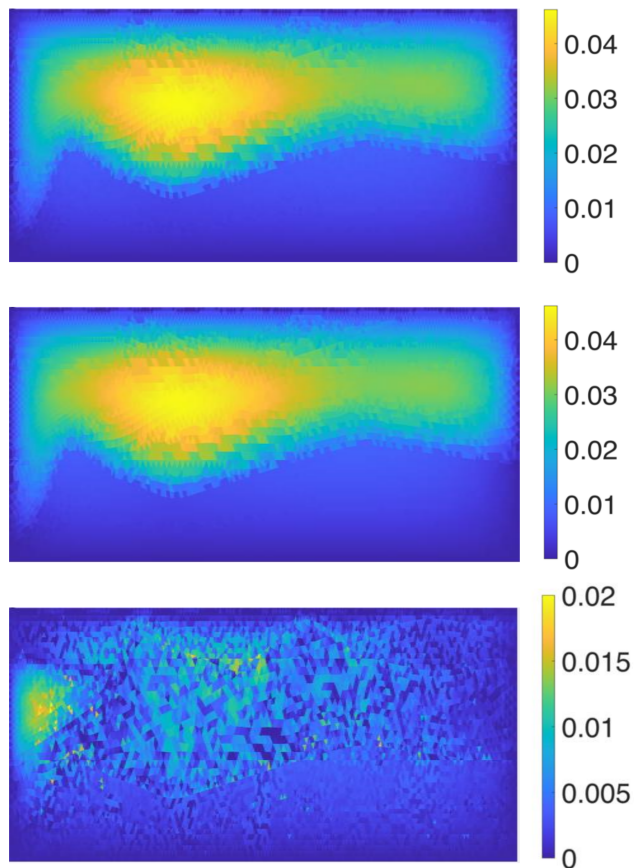


Figure 10: (Top): MLMC-VE approximation of  $\mathbb{E}[u]$  for  $a(\omega, x) = \sum_{r=1}^7 \chi_{D_r}(x) Y_r(\omega)$ , where  $Y_r \sim \mathcal{U}([1, 2])$  for all  $r \neq 7$  and  $Y_7 \sim \mathcal{U}([1, 100])$ . (Middle): MC approximation of  $\mathbb{E}[u]$ . (Bottom): Absolute value of the difference of the two approximations divided by the maximum norm of the MLMC-VE approximation.

smaller interval  $[1, 2]$ ), whereas the seventh region (the red one in Figure 7) has a much larger diffusivity, ranging from 1 to 100. The latter region can therefore be interpreted as a layer made of a different material. The MLMC and the MC estimators, as well as their pointwise absolute difference, are reported in Figure 10. Again, we observe agreement between the two approximations to three decimal digits. Moreover, a comparison between Figures 9 and 10 shows that the pointwise profile of the estimators is affected by the range of variation of the random variables  $\{Y_r\}_{r=1}^7$ , becoming flatter in regions characterized by higher diffusivity. Finally, we consider the QoI defined as the average over the physical domain  $D$  and compare the MC-VE and MLMC-VE estimators of its expected value. We observe agreement to five decimal digits in both configurations: when the diffusivity of all regions is comparable, and when the first six regions have comparable diffusivity, whereas the seventh has a larger diffusivity.

## 8 Conclusions

In this work, we introduced a Monte Carlo Virtual Element method for stochastic elliptic partial differential equations with random diffusion coefficients and proved error estimates for the statistical approximation error of both the solution and relevant (linear) quantities of interest. We also developed and analyzed a Multilevel Monte Carlo Virtual Element method that integrates stochastic sampling with a hierarchy of VE spaces constructed from agglomerated meshes. We proved the convergence of the proposed Multilevel Monte Carlo Virtual Element method and demonstrated that the MLMC-VE strategy yields significant cost reductions compared with the standard Monte Carlo approach. Numerical experiments demonstrate the theoretical bounds and the computational efficiency of the proposed approach. Possible future developments include extending the proposed analysis to the  $p$ -version of the Virtual Element method and stochastic multiphysics models with uncertain data. Another promising direction is the investigation of adaptive multilevel strategies to further reduce computational costs.

## Acknowledgments

The author sincerely thanks Matteo Caldana and Lorenzo Mascotto for their helpful discussions and valuable suggestions during the development of this work.

## References

- [1] P. F. Antonietti, L. Beirão da Veiga, and G. Manzini, editors. *The Virtual Element Method and its Applications*, volume 31 of *SEMA SIMAI Springer Series*. Springer, Cham, Switzerland, 2022.
- [2] P. F. Antonietti, M. Caldana, I. Mazzieri, and A. Re Fraschini. MAGNET: an open-source library for mesh agglomeration by Graph Neural Networks. *Engineering with Computers*, 41(6):4825 – 4850, 2025.
- [3] P. F. Antonietti, M. Corti, and G. Martinelli. Polytopal mesh agglomeration via geometrical deep learning for three-dimensional heterogeneous domains. *Mathematics and Computers in Simulation*, 241:335–353, 2026.
- [4] P. F. Antonietti, N. Farenga, E. Manuzzi, G. Martinelli, and L. Saverio. Agglomeration of polygonal grids using graph neural networks with applications to multigrid solvers. *Computers & Mathematics with Applications*, 154:45–57, 2024.
- [5] A. Barth, C. Schwab, and N. Zollinger. Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numerische Mathematik*, 119(1):123–161, 2011.
- [6] J. Beck, Y. Liu, E. von Schwerin, and R. Tempone. Goal-oriented adaptive finite element multilevel Monte Carlo with convergence rates. *Computer Methods in Applied Mechanics and Engineering*, 402:115582, 2022.
- [7] L. Beirão da Veiga, A. Chernov, L. Mascotto, and A. Russo. Basic principles of *hp* virtual elements on quasiuniform meshes. *Mathematical Models and Methods in Applied Sciences*, 26(8):1567–1598, 2016.
- [8] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of Virtual Element methods. *Mathematical Models and Methods in Applied Sciences*, 23(01):199–214, 2013.
- [9] L. Beirão da Veiga, F. Brezzi, and L. D. Marini. Virtual Elements for linear elasticity problems. *SIAM Journal on Numerical Analysis*, 51(2):794–812, 2013.
- [10] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. The hitchhiker’s guide to the Virtual Element Method. *Mathematical Models and Methods in Applied Sciences*, 24(08):1541–1573, 2014.
- [11] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. The Virtual Element Method. *Acta Numerica*, 32:123–202, 2023.
- [12] S. C. Brenner, Q. Guan, and L.-Y. Sung. Some estimates for virtual element methods. *Computational Methods in Applied Mathematics*, 17(4):553–574, 2017.
- [13] 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 Journal on Numerical Analysis*, 51(1):322–352, 2013.
- [14] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(3), 2011.
- [15] M. Feder, A. Cangiani, and L. Heltai. R3MG: R-tree based agglomeration of polytopal grids with applications to multilevel methods. *Journal of Computational Physics*, 526, 2025.
- [16] M. Giles. Improved Multilevel Monte Carlo Convergence using the Milstein Scheme. In A. Keller, S. Heinrich, and H. Niederreiter, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 343–358, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.
- [17] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.
- [18] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 2015.
- [19] A.-L. Haji-Ali, F. Nobile, E. von Schwerin, and R. Tempone. Optimization of mesh hierarchies in multilevel Monte Carlo samplers. *Stochastics and Partial Differential Equations Analysis and Computations*, 4:76–112, 2016.

- [20] C. Schwab and A. Stein. Multilevel Monte Carlo FEM for elliptic PDEs with Besov random tree priors. *Stochastics and Partial Differential Equations: Analysis and Computations*, 12:1574–1627, 2024.
- [21] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann. Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numerische Mathematik*, 125(3):569–600, 2013.
- [22] O. Čertík, F. Gardini, G. Manzini, L. Mascotto, and G. Vacca. The  $p$ - and  $hp$ -versions of the virtual element method for elliptic eigenvalue problems. *Computers & Mathematics with Applications*, 79(7):2035–2056, 2020.
- [23] Z. Zheng and U. Nackenhorst. Stochastic virtual element methods for uncertainty propagation of stochastic linear elasticity. *Computational Mechanics*, 73(3):667–684, 2024.

## MOX Technical Reports, last issues

Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 31/2026** Guastamacchia, C.; Piersanti, R.; Giardini, F.; Coppini, R.; Ferrantini C.; Dede' L.; Sacconi L.; Regazzoni F.  
*The functional impact of myofiber macroscopic organization and disarray in computational models of the murine heart*
- 30/2026** Regazzoni, F.  
*The internal law of a material can be discovered from its boundary*
- 28/2026** Daniele, F.; Leimer Saglio, C. B.; Pagani, S.; Antonietti, P. F.  
*Mathematical and numerical modeling of coupled oxygen dynamics and neuronal electrophysiology*
- 27/2026** Antonietti, P. F.; Abdalla, O. M. O.; Garroni, M. G.; Mazzieri, I.; Parolini, N.  
*A hybrid reduced-order and high-fidelity discontinuous Galerkin Spectral Element framework for large-scale PMUT array simulations*
- 23/2026** Ballini, E.; Muscarnera, L.; Fumagalli, A.; Scotti, A.; Regazzoni, F.  
*Elimination-compensation pruning for fully-connected neural networks*
- 25/2026** Carrara, D.; Hirschvogel, M.; Bonizzoni, F.; Pagani, S.; Pezzuto, S.; Regazzoni, F.  
*Shape-informed cardiac mechanics surrogates in data-scarce regimes via geometric encoding and generative augmentation*
- 26/2026** Dokuchaev, A.; Bonizzoni, F.; Pagani, S.; Regazzoni, F.; Pezzuto, S.  
*Learning geometry-dependent lead-field operators for forward ECG modeling*
- 21/2026** Bottacini, G.; Torzoni, M.; Manzoni, A.  
*Neural Markov chain Monte Carlo: Bayesian inversion via normalizing flows and variational autoencoders*
- 20/2026** Caldera, L.; Cavinato, L.; Cirone, A.; Cama, I.; Garbarino, S.; Lodi, R.; Tagliavini, F.; Nigri, A.; De Francesco, S.; Cappozzo, A.; Piana, M.; Ieva, F.;  
*DISARM++: Beyond scanner-free harmonization*
- 19/2026** Caldera, L.; Cavinato, L.; Ieva, F.  
*Scanner-agnostic MRI harmonization via SSIM-guided disentanglement*