ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

MATHICSE
Institute of Mathematics
School of Basic Sciences

MATHICSE Technical Report
Nr. 04.2018
March 2018

Analysis of stochastic gradient methods for PDE-constrained optimal control problems with uncertain parameters

Matthieu Martin, Sebastian Krumscheid, Fabio Nobile

https://math.epfl.ch/h

Address: EPFL - SB – INSTITUTE of MATHEMATICS – Math
(Bâtiment MA) Station 8 - CH-1015 - Lausanne - Switzerland

ANALYSIS OF STOCHASTIC GRADIENT METHODS FOR PDE-CONSTRAINED OPTIMAL CONTROL PROBLEMS WITH UNCERTAIN PARAMETERS

M. MARTIN, S. KRUMSCHEID, AND F. NOBILE

Abstract. We consider the numerical approximation of a risk-averse optimal control problem for an elliptic partial differential equation (PDE) with random coefficients. Specifically, the control function is a deterministic, distributed forcing term that minimizes the expected mean squared distance between the state (i.e. solution to the PDE) and a target function, subject to a regularization for well posedness. For the numerical treatment of this risk-averse optimal control problem, we consider a Finite Element discretization of the underlying PDEs, a Monte Carlo sampling method, and gradient type iterations to obtain the approximate optimal control. We provide full error and complexity analysis of the proposed numerical schemes. In particular we compare the complexity of a fixed Monte Carlo gradient method, in which the Finite Element discretization and Monte Carlo sample are chosen initially and kept fixed over the gradient iterations, with a *Stochastic Gradient* method in which the expectation in the computation of the steepest descent direction is approximated by independent Monte Carlo estimators with small sample sizes and possibly varying Finite Element mesh sizes across iterations. We show in particular that the second strategy results in an improved computational complexity. The theoretical error estimates and complexity results are confirmed by our numerical experiments.

Keywords. PDE constrained optimization, risk-averse optimal control, optimization under uncertainty, PDE with random coefficients, stochastic approximation, stochastic gradient, Monte Carlo

AMS subject classifications. 35Q93, 49M99, 65C05, 65N12, 65N30

1. Introduction

Many problems in engineering and science, e.g. shape optimization in aerodynamics or heat transfer in thermal conduction problems, deal with optimization problems constrained by partial differential equations (PDEs) [7, 12, 19, 21, 27]. Often, these types of problems are affected by uncertainties, due to a lack of knowledge, intrinsic variability in the system, or an imprecise manufacturing process. For instance, to determine the optimal cooling of a super-computing center, one should take into account the fact that the heat source from the supercomputers could vary considerably over time and also the heat conduction properties of the machines might not be perfectly determined. As these material properties or boundary conditions are not precisely known, it is reasonable to consider optimal control problems (OCPs) constrained by PDEs with uncertain coefficients, which could be described as random variables or random fields. This OCP is sometimes also referred to as Optimization Under Uncertainty (OUU).

In this work we focus on the numerical approximation of the problem of controlling the solution of an elliptic PDE with random coefficients by a distributed unconstrained control. Specifically, the control acts as a volumetric forcing term, so that the solution is as close as possible to a given target function.

Date: March 9, 2018.

While there is a vast literature on the numerical approximation of PDE-constrained optimal control problems (see e.g. [7, 21] and references therein) in the deterministic case, as well as on the numerical approximation of (uncontrolled) PDEs with random coefficients (see e.g. [3, 18, 29] and references therein), the analysis of corresponding PDE constrained control problem under uncertainty is much more recent and incomplete, although the topic has received increasing attention in the last few years.

The formulations of the PDE-constrained OCPs under uncertainty that can be found in the literature can be roughly grouped in two categories.

In the first category, the control is *random* [1, 6, 10, 25, 35, 40]. This situation arises when the randomness in the PDE is observable hence an optimal control can be built for each realization of the random system. However, the corresponding optimality system might still be fully coupled in the random parameters if the objective function involves some statistics of the state variables. The dependence on the random parameters is typically approximated either by polynomial chaos expansions or Monte Carlo (MC) techniques.

The former approach is considered e.g. in [25], where the authors prove analytic dependence of the control on the random parameters and study its best N-term polynomial chaos approximation for a linear parabolic PDE-constrained OCP; the work [10], combines a stochastic collocation with a Finite Element (FE) based reduced basis method to alleviate the computational effort; the works [6, 35, 40] address the case of a fully coupled optimality system discretized by either Galerkin or collocation approaches and propose different methods, such as sequential quadratic programming, or block diagonal preconditioning to solve the coupled system efficiently. Monte Carlo and Multilevel Monte Carlo approaches are considered in [1] instead, where the case of random coefficients with limited spatial regularity is addressed.

In the second category, the control is *deterministic* [2, 9, 17, 22, 23, 24, 41]. This situation arises when randomness in the system is not observable at the time of designing the control, so that the latter should be *robust* in the sense that it minimizes the *risk* of obtaining a solution which leads to high values of the objective function. This situation is also referred to as *risk-averse optimal control* and always leads to a fully coupled optimality system in the random parameters. The idea of minimizing a risk to obtain a solution with favorable properties goes back to the origins of robust optimization [39]. Here, *risk* refers to a suitable statistical measure of the objective function to be minimized, such as its expectation, expectation plus variance, a quantile, or a conditional expectation above a quantile (so called Conditional Value at Risk (CVaR) [34]).

Numerical methods for OCPs of this category typically depend on the choice of the risk measure. For example, the work [2] considers a risk measure that involves the mean and variance of the objective function and uses second order Taylor expansions combined with randomized estimators to reduce the computational effort. The work [41] considers a risk measure that involves only the mean of the objective function (hereafter named mean-based risk), with an additional penalty on the variance of the state, and proposes a gradient type method, in which the expectation of the gradient is computed by a Multilevel Monte Carlo method. In [9], the authors also consider a mean-based risk problem and propose a reduced basis method on the space of controls to dramatically reduce the computational effort. In the work [22], the author presents a more general type of OCP, using the general notion of a risk measure, and derives the corresponding optimality system of PDEs to be solved. For its numerical solution, a trust-region Newton conjugate gradient algorithm is proposed in [23], combined with an adaptive sparse grid collocation for the discretization of the PDE in the stochastic space. The work [24] considers derivative-based optimization methods for the robust CVaR risk measure, which are building upon introducing smooth approximations to the CVaR. Finally, in the work [17], the authors consider a boundary OCP where the deterministic control appears as a Neumann boundary condition.

In this work, we follow the second modeling category consider the (robust) OCP of minimizing the mean-based risk of the objective function. We consider in particular gradient type methods where adjoint calculus is used to represent the gradient of the objective function, and FE approximations of the primal and dual problems, as well as a Monte Carlo approximation of the expectation in the risk measure are employed. The reason for looking at Monte Carlo approximations, instead of polynomial chaos ones, is to develop methods that can potentially handle many random parameters and possibly rough random coefficients.

Our main contribution is to provide a full error analysis including the finite element, the Monte Carlo and the gradient iterations errors, as well as a complexity analysis when all sources of errors are optimally balanced to achieve a given tolerance. The motivation for analyzing gradient type optimization methods is twofold. First, their rather simple structure allows for a complete complexity analysis, which is desirable in practice due to their wide-spread use. Second, our analysis reveals that the cost due to the FE and the Monte Carlo approximations dominate the overall computational complexity, in the sense that the gradient type method only increases the cost by a logarithmic term.

It is noteworthy that other error analysis have been proposed in [10] in the case of a random control, with a discretization in space by Finite Elements and in probability by stochastic collocation, and in [17] in the case of a mean-based risk for a deterministic boundary control problem, using a Finite Element discretization both in space and in probability.

The first gradient method that we consider is the standard gradient method (which we call fixed MC gradient), in which the Finite Element discretization and the Monte Carlo sample are chosen initially and kept fixed over the iterations of the gradient method. If N is the sample size of the Monte Carlo estimator, this method entails the solution of N primal and N dual problems at each iteration of the gradient method, which could be troublesome if a small tolerance is required, entailing a very large N and small Finite Element mesh size.

We then turn to stochastic versions of the gradient method in which the gradient is re-sampled independently at each iteration and the Finite Element mesh size can be refined along the iterations. This corresponds to taking, at each iteration, an independent Monte Carlo estimator with only one realization ($N = 1$) or a very small and fixed sample size ($N = \bar{N}$) independently of the required tolerance, with possibly a finer Finite Element mesh. We follow, in particular, the Robbins-Monroe strategy [30, 33, 36] of reducing progressively the step-size to achieve convergence of the Stochastic Gradient iterations.

Stochastic Gradient (SG) techniques have been extensively applied to machine learning problems [13, 14, 16, 26], but have not yet been used for risk-averse PDE-constrained optimization problems. Here, we show that our Stochastic Gradient method improves the complexity of the fixed MC gradient method by a logarithmic factor. Although the computational gain is not dramatic, we see potential in this approach as only one primal problem and one dual problem have to be solved at every iteration of the gradient method. Moreover, we believe that the whole construction is more amenable to an adaptive version, which, in combination with an appropriate error estimator, allows for a self-controlling algorithm. We leave this for future work.

The rest of the paper is organized as follows: in Section 2 we set the mean-based risk-averse optimal control problem and recall its well posedness and the optimality conditions; in Sections 3, 4, 5 we introduce, respectively, the finite element discretization, the Monte Carlo approximation, and the steepest descent (gradient) method, including their full error analysis. In particular, Theorem 5 in Section 5 gives an error bound for the fully discrete solution of the fixed MC gradient method, whereas Corollary 2 gives the corresponding computational complexity. In Section 6 we analyze the Stochastic Gradient method with fixed finite element discretization over the iterations (with error bound given in Theorem 6 and the corresponding complexity result in Corollary 3), whereas in Section 7 we analyze the Stochastic Gradient version in which the Finite Element mesh is refined over the iterations (Theorem 8 and Corollary 4). In Section 8, we discuss a 2D test problem and confirm numerically the theoretical error bounds and complexities derived in the preceding Sections. Finally, in Section 9 we draw some conclusions.

2. Problem setting

We start introducing the primal problem that will be part of the OCP discussed in the following. Specifically, we consider the problem of finding the solution $y : D \times \Gamma \to \mathbb{R}$ of the elliptic random PDE

$$(1)\qquad \begin{cases} -\operatorname{div}(a(x,\omega)\nabla y(x,\omega)) &= \phi(x,\omega), & x \in D, \quad \omega \in \Gamma, \\ y(x,\omega) &= 0, & x \in \partial D, \quad \omega \in \Gamma, \end{cases}$$

where $D \subset \mathbb{R}^n$ is open and bounded, denoting the physical domain, $(\Gamma, \mathcal{F}, P)$ is a complete probability space, and $\omega \in \Gamma$ is an elementary random event. The diffusion coefficient a is an almost surely (a.s.) continuous and positive random field on D, and ϕ is a stochastic source term (that could contain, for example, a deterministic control part).

Before addressing the optimal control problem related to the random PDE (1), we will first recall the well posedness results for (1). We begin by recalling some usual functional spaces needed for the analysis that follows. Let $L^p(D)$ for $1 \le p < \infty$ denote the space of functions for which the p-th power of their absolute value is Lebesgue integrable, that is

$$L^p(D) = \{y : D \to \mathbb{R}, \;\; f \text{ measurable}, \;\; \text{and} \;\; \int_D |y|^p \mathrm{d}x < +\infty\},$$

and $L^\infty(D)$ the space of measurable functions that are bounded almost everywhere (a.e.) on D. Throughout this work, we will denote by $\|\cdot\| \equiv \|\cdot\|_{L^2(D)}$ the usual $L^2(D)$-norm induced by the inner product $\langle f, g\rangle = \int_D fg\mathrm{d}x$ for any $f, g \in L^2(D)$. Furthermore, we introduce the Sobolev spaces

$$H^1(D) = \{y \in L^2(D), \quad \partial_{x_i} y \in L^2(D), \quad i = 1, \dots, n\}$$

and

$$H^1_0(D) = \{y \in H^1(D), \quad y|_{\partial D} = 0\}.$$

We use the equivalent H^1-norm on the space $H^1_0(D)$ defined by $\|y\|_{H^1(D)} = \|y\|_{H^1_0(D)} = \|\nabla y\|$ for any $y \in H^1_0(D)$. Moreover, we recall the Poincaré inequality for any function $y \in H^1_0(D)$

$$\|y\| \le C_p \|\nabla y\| = C_p \|y\|_{H^1(D)},$$

where C_p is the Poincaré constant, and that $H^{-1}(D) = \left(H^1_0(D)\right)^*$ is the topological dual of $H^1_0(D)$. For $r \in \mathbb{N}$ we further recall the space $H^r(D)$ of $L^2(D)$ functions with all partial derivatives up to order r in $L^2(D)$ with norm $\|y\|_{H^r(D)}$ and semi-norm $|y|_{H^r(D)}$ given by

$$\|y\|^2_{H^r(D)} = \sum_{|\boldsymbol{\alpha}|\leq r} \left\| \frac{\partial^{|\boldsymbol{\alpha}|} y}{\partial x^{\boldsymbol{\alpha}}} \right\|^2_{L^2(D)} \quad \text{and} \quad |y|^2_{H^r(D)} = \sum_{|\boldsymbol{\alpha}|=r} \left\| \frac{\partial^{|\boldsymbol{\alpha}|} y}{\partial x^{\boldsymbol{\alpha}}} \right\|^2_{L^2(D)},$$

respectively, for the multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$. Finally, we introduce the Bochner spaces $L^p(\Gamma, \mathcal{V})$, which are formal extensions of Lebesgue spaces $L^p(\Gamma)$, for functions with values in a separable Hilbert space $\mathcal{V}$ as

$$L^p(\Gamma, \mathcal{V}) = \{y : \Gamma \to \mathcal{V},\ \ y \text{ measurable}, \int_\Gamma \|y(\omega)\|^p_{\mathcal{V}} \mathrm{d}P(\omega) < +\infty\},$$

equipped with the norm $\|y\|_{L^p(\Gamma,\mathcal{V})} = \left(\int_\Gamma \|y(\omega)\|^p_{\mathcal{V}} \mathrm{d}P(\omega)\right)^{\frac{1}{p}}$, see, e.g., [15] for details.

As it is common for the well posedness of the elliptic PDE (1), we assume that the diffusion coefficient a in (1) is uniformly elliptic.

Assumption 1. *The diffusion coefficient $a \in L^\infty(D \times \Gamma)$ is bounded and bounded away from zero a.e. in $D \times \Gamma$, i.e.*

$$\exists \quad a_{\min}, a_{\max} \in \mathbb{R} \quad \text{such that} \quad 0 < a_{\min} \leq a(x,\omega) \leq a_{\max} \quad \text{a.e. in } D \times \Gamma.$$

Now we are in the position to recall the well posedness of the random PDE (1), which is a standard result, see e.g. [4, 29].

Lemma 1 (Well posedness of (1))**.** *Let Assumption 1 hold. If $\phi \in L^2(\Gamma, H^{-1}(D))$, then problem* (1) *admits a unique solution $y \in L^2(\Gamma, H^1_0(D))$ s.t.*

$$\|y(\cdot,\omega)\|_{H^1_0(D)} \leq \frac{1}{a_{min}} \|\phi(\cdot,\omega)\|_{H^{-1}(D)} \quad \text{for a.e. } \omega \in \Gamma$$

$$\text{and} \ \ \|y\|_{L^2(\Gamma,H^1_0(D))} \leq \frac{1}{a_{min}} \|\phi\|_{L^2(\Gamma,H^{-1}(D))}.$$

Finally, as we will occasionally need H^2-regularity in the following Sections, we also introduce a sufficient condition on the domain D and on the gradient of a.

Assumption 2. *The domain $D \subset \mathbb{R}^n$ is polygonal convex and the random field $a \in L^\infty(D \times \Gamma)$ is such that $\nabla a \in L^\infty(D \times \Gamma)$,*

Then, using standard regularity arguments for elliptic PDEs, one can prove the following result [15].

Lemma 2. *Let Assumptions 1 and 2 hold. If $\phi \in L^2(\Gamma, L^2(D))$, then problem* (1) *has a unique solution $y \in L^2(\Gamma, H^2(D))$. Moreover there exists a constant C, independent of ϕ, such that*

$$\|y\|_{L^2(\Gamma,H^2(D))} \leq C\|\phi\|_{L^2(\Gamma,L^2(D))}.$$

We are now ready to introduce the optimal control problem linked with the PDE (1), which we will study in the rest of the paper.

2.1. **Optimal Control Problem.** We define the primal problem for the OCP as the elliptic PDE (1), by particularizing its right hand side to:

$$\begin{cases} -\operatorname{div}(a(x,\omega)\nabla y(x,\omega)) &=& g(x) + u(x), & x \in D, \quad \omega \in \Gamma, \\ y(x,\omega) &=& 0, & x \in \partial D, \quad \omega \in \Gamma, \end{cases} \tag{2}$$

with $g \in H^{-1}(D)$ and $u \in U$, where $U \subset L^2(D)$ denotes the set of all admissible (deterministic) control functions. We set the state space of the solution to (2) as $Y = H^1_0(D)$. To emphasize the dependence of the solution to the PDE on the control function and on a particular realization $a(\cdot,\omega)$ of the random field, we will use the notation $y_\omega(u)$. When the particular realization of a is not relevant, or when no confusion arises, we will also simply write $y(u)$ from times. In this work, we focus on the objective function

$$(3) \qquad J(u) = \mathbb{E}[f(u,\omega)] \quad \text{with} \quad f(u,\omega) = \frac{1}{2}\|y_\omega(u) - z_d\|^2 + \frac{\alpha}{2}\|u\|^2,$$

where z_d is a given target function which we would like the state $y_\omega(u)$ to approach as close as possible, in a mean-square-error sense. The coefficient $\alpha \geq 0$ is a constant of the problem that models the price of energy, i.e. how expensive it is to add some energy in the control u in order to decrease the first distance term $\mathbb{E}\left[\|y_\omega(u) - z_d\|^2\right]$. The ultimate goal then is the OCP, of determining the optimal control u^*, so that

$$(4) \qquad u^* \in \operatorname*{arg\,min}_{u\in U} J(u), \quad \text{s.t.} \quad y_\omega(u) \in Y \quad \text{solves} \quad (2) \quad \text{a.s.}$$

Remark 1. *The optimal control u^* in* (4) *is the one that provides the best fit* $\|y_\omega(u^*) - z_d\|$ on average *not requiring too much* control energy *(induced by the regularization term). In view of applications, one may consider a more general objective function $J(u) = \sigma\Big(\frac{1}{2}\|y_\omega(u) - z_d\|^2\Big) + \frac{\alpha}{2}\|u\|^2$, where $\sigma(\cdot)$ is a* more robust risk measure *such as the Conditional Value at Risk [24]. In this paper, however, we restrict to the simple* expectation *risk measure, namely $\sigma(\cdot) = \mathbb{E}[\cdot]$, for sake of simplicity.*

As we aim at minimizing the functional J, we will use the theory of optimization and calculus of variations. Specifically, we introduce the optimality condition for the OCP (4), in the sense that the optimal control u^* satisfies

$$(5) \qquad \langle \nabla J(u^*), v - u^*\rangle \geq 0 \quad \forall v \in Y.$$

Here, by $\nabla J(u)$ we denote the $L^2(D)$-functional representation of the Gateaux derivative of J, namely

$$\int_D \nabla J(u)\delta u \,\mathrm{d}x = \lim_{\epsilon\to 0} \frac{J(u+\epsilon\delta u) - J(u)}{\epsilon} = DJ(u)(\delta u) \quad \forall\, \delta u \in L^2(D).$$

In order to study the well posedness of problem (4), we introduce a further assumption on α, U and g.

Assumption 3. *The regularization parameter α is strictly positive, i.e. $\alpha > 0$. Moreover, the space of admissible control functions is $U = L^2(D)$ and the deterministic source term is such that $g \in L^2(D)$.*

It follows from the results in [22] that problem (4) is well posed. As a matter of fact, problem (4) is even well posed for more general settings than the one considered here. For completeness, we give a short proof for the particular setting considered in this work, as many of the following results will build on it. For this we first introduce the following solution operator corresponding to the elliptic PDE (1):

$$\begin{aligned} S : L^2(\Gamma, H^{-1}(D)) &\longrightarrow L^2(\Gamma, Y) \\ \phi &\longmapsto S\phi = y \text{ solution of } (1). \end{aligned}$$

Notice that the operator S is continuous in view of Lemma 1. In the case of $\phi = g + u \in L^2(D)$ deterministic, we will sometimes use the notation $S_\omega(g+u) = y_\omega(u)$ to denote one ω-realization of y. As S is self-adjoint, we have $S^* = S$. Moreover, for any separable Hilbert space $\mathcal{V}$, we denote by $\mathbb{E}$ the usual expectation operator with respect to (w.r.t.) the probability measure P acting on the space $L^2(\Gamma, \mathcal{V})$, i.e. $\mathbb{E} : L^2(\Gamma, \mathcal{V}) \to \mathcal{V}$. Its adjoint operator is

$$\begin{aligned} \mathbb{E}^* : \mathcal{V}' &\longrightarrow L^2(\Gamma, \mathcal{V}') \\ v &\longmapsto v, \end{aligned}$$

which associates the *constant* stochastic (i.e. deterministic) function $v \in L^2(\Gamma, \mathcal{V}')$ to each deterministic function $v \in \mathcal{V}'$. Finally we define the two operators

$$\widetilde{S} = S\mathbb{E}^* : L^2(D) \to L^2(\Gamma, Y) \quad \text{and} \quad \widetilde{S}^* = \mathbb{E}S^* : L^2(\Gamma, Y) \to L^2(D).$$

Existence and uniqueness of the OCP (4) can then be stated as follows.

Theorem 1. *Suppose Assumptions 1 and 3 hold. Then the OCP* (4) *admits a unique control $u^* \in U$. Moreover*

$$\nabla J(u) = \alpha u + \mathbb{E}[p_\omega(u)], \tag{6}$$

where $p_\omega(u) = p$ is the solution of the adjoint problem (a.s. in Γ)

$$\begin{cases} -\operatorname{div}(a(\cdot,\omega)\nabla p(\cdot,\omega)) &= y(\cdot,\omega) - z_d & \text{in } D, \\ p(\cdot,\omega) &= 0 & \text{on } \partial D. \end{cases} \tag{7}$$

Proof. Let us define the inner product $\ll \cdot,\cdot \gg$ on the Bochner space $L^2(\Gamma, U)$, $\ll u, v \gg = \mathbb{E}\left[< u, v >\right] = \int_\Gamma \int_D u(x,\omega)v(x,\omega)\mathrm{d}x\, \mathrm{d}P(\omega)$. Using the linearity of the introduced operators, we can write $J(u)$ as

$$\begin{aligned} J(u) &= \frac{1}{2}\mathbb{E}\left[< \widetilde{S}(g+u) - z_d, \widetilde{S}(g+u) - z_d >\right] + \frac{\alpha}{2} < u, u > \\ &= \frac{1}{2} \ll \widetilde{S}(g+u) - z_d, \widetilde{S}(g+u) - z_d \gg + \frac{\alpha}{2} < u, u > \\ &= \frac{1}{2} \ll \widetilde{S}u, \widetilde{S}u \gg + \ll \widetilde{S}g - z_d, \widetilde{S}u \gg + \frac{1}{2} \ll \widetilde{S}g - z_d, \widetilde{S}g - z_d \gg + \frac{\alpha}{2} < u, u > . \end{aligned}$$

Defining the bi-linear form $A : U \times U \to \mathbb{R}$, $A(u, v) = \ll \widetilde{S}u, \widetilde{S}v \gg + \alpha < u, v >$, the linear form $G : U \to \mathbb{R}$, $G(v) = \ll \widetilde{S}g - z_d, \widetilde{S}v \gg$, and the constant $k = \frac{1}{2} \ll \widetilde{S}g - z_d, \widetilde{S}g - z_d \gg \in \mathbb{R}$, we find

$$J(u) = \frac{1}{2}A(u, u) + G(u) + k.$$

Thanks to Assumptions 1 and 3, it is easy to see that A is coercive and continuous (cf. Lemma 1), that G is continuous, and that $k < +\infty$. Then, applying Thm. 7.1 of [28], we conclude that there exists a unique solution $u^* \in U$ to problem (4). Next, we compute the Gâteaux derivative of J at the point u in the direction δu:

$$\begin{aligned} DJ(u)(\delta u) = \int_D \nabla J(u)\delta u \mathrm{d}x &= A(u, \delta u) + G(\delta u) \\ &= \ll \widetilde{S}u, \widetilde{S}\delta u \gg + \alpha < u, \delta u > + \ll \widetilde{S}g - z_d, \widetilde{S}\delta u \gg \\ &= < \widetilde{S}^*(\widetilde{S}(g+u) - z_d), \delta u > + < \alpha u, \delta u > \\ &= < \mathbb{E}S^*(S(g+u) - z_d), \delta u > + < \alpha u, \delta u > \\ &= < \alpha u + \mathbb{E}\left[S^*(S(g+u) - z_d)\right], \delta u > . \end{aligned}$$

Defining $p_\omega(u)$ as $p_\omega(u) = S^*\left(S(g+u) - z_d\right) = S^*\left(y_\omega(u) - z_d\right)$, which is the solution of equation (7), we get $\nabla J(u) = \alpha u + \mathbb{E}\left[p_\omega(u)\right]$. □

Remark 2. *By computing the gradient of f w.r.t. u, we can easily get $\nabla f(u,\omega) = \alpha u + p_\omega(u)$. Consequently, the previous proof, also reveals that*

$$\nabla J(u) = \nabla\mathbb{E}[f(u,\omega)] = \mathbb{E}\left[\nabla f(u,\omega)\right].$$

In Theorem 1, $p_\omega(u) = p$ is the so-called adjoint function associated to the elliptic PDE (2) and satisfies the adjoint equation which depends on the solution $y = y_\omega(u)$. As p depends on u through y, we will also write $p(y_\omega(u))$ for $p_\omega(u)$ from times.

For notational convenience, we introduce the weak formulation of (2), which reads

$$\text{find } y_\omega \in Y \ s.t. \ b_\omega(y_\omega, v) = \langle g + u, v\rangle \quad \forall v \in Y \qquad \text{for a.e. } \omega \in \Gamma, \tag{8}$$

where $b_\omega(y, v) := \int_D a(\cdot, \omega)\nabla y \nabla v dx$. Similarly, the weak form of problem (7) reads:

$$b_\omega(v, p_\omega) = \langle v, y_\omega - z_d\rangle \quad \forall v \in Y \qquad \text{for a.e. } \omega \in \Gamma. \tag{9}$$

We can thus rewrite the OCP (4) equivalently as:

$$\begin{cases} \min_{u\in U} J(u) = \frac{1}{2}\mathbb{E}[\|y_\omega(u) - z_d\|^2] + \frac{\alpha}{2}\|u\|^2 \\ \text{s.t.} \quad y_\omega \in Y \quad \text{solving} \\ b_\omega(y_\omega, v) = \langle g + u, v\rangle \quad \forall v \in Y \qquad \text{for a.e. } \omega \in \Gamma. \end{cases} \tag{10}$$

As we want to compute numerically the problem solution, we introduce in the following Section a Finite Element approximation and different version of error estimates.

3. Finite Element approximation in physical space

In this section we analyze the semi-discrete OCP obtained by approximating the underlying PDE by a Finite Element method. In particular, we provide a priori error bounds for the optimal control. Let us denote by $\{\tau_h\}_{h>0}$ a family of regular triangulations of D. Furthermore, let Y^h be the space of continuous piece-wise polynomial functions of degree r over τ_h that vanish on ∂D, i.e. $Y^h = \{y \in C^0(\overline{D}) : y|_K \in \mathbb{P}_r(K) \quad \forall K \in \tau_h, y|_{\partial D} = 0\} \subset Y = H^1_0(D)$. Finally, we set $U^h = Y^h$. We can then reformulate the OCP (10) as a finite dimensional OCP in the FE space:

$$\begin{cases} \min_{u^h\in U^h} J^h(u^h) := \frac{1}{2}\mathbb{E}[\|y^h_\omega(u^h) - z_d\|^2] + \frac{\alpha}{2}\|u^h\|^2 \\ \text{s.t. } y^h_\omega \in Y^h \text{ and} \\ b_\omega(y^h_\omega(u^h), v^h) = \langle u^h + g, v^h\rangle \quad \forall v^h \in Y^h \quad \text{for a.e. } \omega \in \Gamma. \end{cases} \tag{11}$$

Analogously to the (continuous) solution operator S of (1) introduced in Section 2.1, here we introduce its discrete version associated to problem (11). That is, let $S^h_\omega : U \to Y^h$ be such that $y^h_\omega = S^h_\omega(g + u^h)$ solves $b_\omega(y^h_\omega, v^h) = \langle g + u^h, v^h\rangle \quad \forall v^h \in Y^h$. We also introduce the L^2-projection operator onto U^h, denoted by $g^h = \Pi_{U^h}(g)$, as

$$\forall q \in U, \ \ \langle \Pi_{U^h} q, v^h\rangle = \langle q, v^h\rangle \ \ \forall v^h \in U^h.$$

As mentioned before, we may suppress the index ω of S_ω when no ambiguity arises, we do so also for $S^h_\omega = S^h$. Moreover, we denote by $\left(S^h\right)^*$ the corresponding adjoint operator of S^h. From now on, and throughout the rest of this paper, we assume that Assumptions 1, 2 and 3 are verified. Then we can state the following FE approximation result.

Lemma 3. *The discrete OCP* (11) *is well posed and ∇J^h can be characterized as*

$$\nabla J^h(u^{h*}) = \Pi_{U^h}(\alpha u^{h*} + \mathbb{E}[p^h(u^{h*})]) \tag{12}$$

and

$$p^h(u^{h*}) := \left(S^h\right)^* (S^h(u^{h*} + g) - z_d) \in L^2(\Gamma, Y^h).$$

Remark 3. *Notice that since we defined $U^h = Y^h$, it follows that $\mathbb{E}[p^h(u^{h*})] \in U^h$ and $\nabla J^h(u^{h*}) = \alpha u^{h*} + \mathbb{E}[p^h(u^{h*})]$.*

Following similar arguments as in Thm. 3.4 of [21] and using the optimality condition, and the weak form of the primal and dual problems, we can prove the following.

Theorem 2. *Let u^* be the optimal control solution of problem* (10) *and denote by u^{h*} the solution of the approximated problem* (11)*. Then it holds that*

$$\frac{\alpha}{2}\|u^*-u^{h*}\|^2+\frac{1}{2}\mathbb{E}[\|y(u^*)-y^h(u^{h*})\|^2] \leq \frac{1}{2\alpha}\mathbb{E}[\|p(u^*)-\widetilde{p}^h(u^*)\|^2]+\frac{1}{2}\mathbb{E}[\|y(u^*)-y^h(u^*)\|^2], \tag{13}$$

where, $\widetilde{p}^h(u^*)=\widetilde{p}^h_\omega(u^*)$ is such that

$$b_\omega(v^h,\widetilde{p}^h_\omega)=\langle v^h, y_\omega - z_d\rangle \quad \forall v^h\in Y^h \ \text{ for a.e. } \omega\in\Gamma. \tag{14}$$

Proof. It follows from Theorem 1 and Lemma 3 that the FE version of the optimality condition (5) reads:

$$\langle \nabla J^h(u^{h*}), v^h-u^{h*}\rangle \geq 0 \quad \forall v^h\in U^h. \tag{15}$$

Choosing $v=u^{h*}\in Y^h\subset Y$ in (5), $v^h=\Pi_{U^h}(u^*)$ in (15), and observing that

$$0\leq \langle\nabla J^h(u^{h*}),\Pi_{U^h}(u^*)-u^{h*}\rangle = \langle\nabla J^h(u^{h*}),\Pi_{U^h}(u^*-u^{h*})\rangle = \langle\nabla J^h(u^{h*}),u^*-u^{h*}\rangle,$$

since $\nabla J^h(u^{h*})\in U^h$, we obtain

$$\langle\alpha(u^*-u^{h*})+\mathbb{E}[p(u^*)]-\mathbb{E}[p^h(u^{h*})],u^{h*}-u^*\rangle\geq 0.$$

Then introducing $\widetilde{p}^h(u^*)=\left(S^h\right)^*(S(u^*+g)-z_d)$, which belongs to $L^2(\Gamma,Y^h)$ since the two operators S and $\left(S^h\right)^*$ are bounded, we obtain

$$\alpha\|u^*-u^{h*}\|^2\leq\langle\mathbb{E}[p(u^*)]-\mathbb{E}[\widetilde{p}^h(u^*)]+\mathbb{E}[\widetilde{p}^h(u^*)]-\mathbb{E}[p^h(u^{h*})],u^{h*}-u^*\rangle. \tag{16}$$

In the following, we will repeatedly use the primal and dual weak formulations (8),(9) and(14), for the continuous problem and its FE approximation, yielding

$$\begin{aligned}
\langle\widetilde{p}^h_\omega(u^*)-p^h_\omega(u^{h*}),u^{h*}-u^*\rangle &= b_\omega(y^h_\omega(u^{h*})-y^h_\omega(u^*),\widetilde{p}^h_\omega(u^*)-p^h_\omega(u^{h*}))\\
&=\int_D \underbrace{(y^h_\omega(u^{h*})-y^h_\omega(u^*))}_{\pm y_\omega(u^*)}(y_\omega(u^*)-y^h_\omega(u^{h*}))\mathrm{d}x\\
&=-\|y_\omega(u^*)-y^h_\omega(u^{h*})\|^2+\int_D(y_\omega(u^*)-y^h_\omega(u^*))(y_\omega(u^*)-y^h_\omega(u^{h*}))\mathrm{d}x\\
&\leq-\|y_\omega(u^*)-y^h_\omega(u^{h*})\|^2+\frac{1}{2}\|y_\omega(u^*)-y^h_\omega(u^*)\|^2+\frac{1}{2}\|y_\omega(u^*)-y^h_\omega(u^{h*})\|^2\\
&\leq-\frac{1}{2}\|y_\omega(u^*)-y^h_\omega(u^{h*})\|^2+\frac{1}{2}\|y_\omega(u^*)-y^h_\omega(u^*)\|^2.
\end{aligned}$$

Taking the mean over all realizations $\omega\in\Gamma$, using (16), and Fubini's theorem we have that

$$\begin{aligned}
\alpha\|u^*-u^{h*}\|^2+\frac{1}{2}\mathbb{E}[\|y(u^*)-y^h(u^{h*})\|^2] &\leq \mathbb{E}[\langle p(u^*)-\widetilde{p}^h(u^*),u^{h*}-u^*\rangle]+\frac{1}{2}\mathbb{E}[\|y(u^*)-y^h(u^*)\|^2]\\
&\leq\frac{1}{2\alpha}\|p(u^*)-\widetilde{p}^h(u^*)\|^2+\frac{\alpha}{2}\|u^{h*}-u^*\|^2+\frac{1}{2}\mathbb{E}[\|y(u^*)-y^h(u^*)\|^2],
\end{aligned}$$

which leads to the claim. □

The FE error $\|u^*-u^{h*}\|$ is thus completely determined by the approximation properties of the discrete solution operators S^h and $\left(S^h\right)^*$. Using similar arguments as in [21, Thm. 3.5], we can also control the FE error of the state variable in H^1, i.e. of $\|y(u^*)-y^h(u^{h*})\|_{H^1_0}$.

Theorem 3. *With the same notations as in Theorem 2, there exists a constant $C > 0$ independent of h such that*

$$\begin{aligned}(17)\quad &\|u^* - u^{h*}\|^2 + \mathbb{E}[\|y(u^*) - y^h(u^{h*})\|^2] + h^2\mathbb{E}[\|y(u^*) - y^h(u^{h*})\|^2_{H_0^1}] \\ &\le C\{\mathbb{E}[\|p(u^*) - \widetilde{p}^h(u^*)\|^2] + \mathbb{E}[\|y(u^*) - y^h(u^*)\|^2] + h^2\mathbb{E}[\|y(u^*) - y^h(u^*)\|^2_{H_0^1}]\}.\end{aligned}$$

Proof. From the uniform coercivity of the bi-linear form $b_\omega(\cdot,\cdot)$, c.f. Assumption 1, it immediately follows

$$\|y_\omega - y_\omega^h\|^2_{H_0^1} \le \frac{1}{a_{min}}\left\{b_\omega\big(y_\omega - y_\omega^h, y_\omega - \widetilde{y}_\omega^h\big) + b_\omega\big(y_\omega - y_\omega^h, \widetilde{y}_\omega^h - y_\omega^h\big)\right\},$$

where we have used the notation $y_\omega = y_\omega(u^*)$, $y_\omega^h = y_\omega^h(u^{h*})$, and $\widetilde{y}_\omega^h = y_\omega^h(u^*)$. Moreover

$$\begin{aligned}\frac{1}{a_{min}}b_\omega\big(y_\omega - y_\omega^h, y_\omega - \widetilde{y}_\omega^h\big) &\le \frac{a_{max}}{a_{min}}\|y_\omega - y_\omega^h\|_{H_0^1}\|y_\omega - \widetilde{y}_\omega^h\|_{H_0^1} \\ &\le \frac{1}{4}\|y_\omega - y_\omega^h\|^2_{H_0^1} + \frac{a_{max}^2}{a_{min}^2}\|y_\omega - \widetilde{y}_\omega^h\|^2_{H_0^1},\end{aligned}$$

as well as

$$\begin{aligned}\frac{1}{a_{min}}b_\omega\big(y_\omega - y_\omega^h, \widetilde{y}_\omega^h - y_\omega^h\big) &\le \frac{1}{a_{min}}\langle u^* - u^{h*}, \widetilde{y}_\omega^h - y_\omega^h\rangle \\ &\le \frac{1}{a_{min}}\langle u^* - u^{h*}, \widetilde{y}_\omega^h - y_\omega\rangle + \frac{1}{a_{min}}\langle u^* - u^{h*}, y_\omega - y_\omega^h\rangle \\ &\le \frac{C_p^2}{2a_{min}}\|u^* - u^{h*}\|^2 + \frac{1}{2a_{min}}\|y_\omega - \widetilde{y}_\omega^h\|^2_{H_0^1} + \frac{C_p^2}{a_{min}^2}\|u^* - u^{h*}\|^2 + \frac{1}{4}\|y_\omega - y_\omega^h\|^2_{H_0^1}.\end{aligned}$$

Finally, it follows that

$$\|y_\omega - y_\omega^h\|^2_{H_0^1} \le C\{\|y_\omega - \widetilde{y}_\omega^h\|^2_{H_0^1} + \|u^* - u^{h*}\|^2\}$$

and

$$h^2\mathbb{E}[\|y_\omega - y_\omega^h\|^2_{H_0^1}] \le h^2 C\{\mathbb{E}[\|y_\omega - \widetilde{y}_\omega^h\|^2_{H_0^1}] + \|u^* - u^{h*}\|^2\},$$

which, combined with (13), completes the proof. □

We can now proceed and estimate the right hand side of (17), assuming the primal and dual solutions are sufficiently smooth.

Corollary 1. *Suppose that $y(u^*), p(u^*) \in L^2(\Gamma, H^{r+1}(D))$, then we have*

$$\begin{aligned}(18)\quad &\|u^* - u^{h*}\|^2 + \mathbb{E}[\|y(u^*) - y^h(u^{h*})\|^2] + h^2\mathbb{E}[\|y(u^*) - y^h(u^{h*})\|^2_{H_0^1}] \\ &\qquad\le Ch^{2r+2}\{\mathbb{E}[|y_\omega(u^*)|^2_{H^{r+1}}] + \mathbb{E}[|p_\omega(u^*)|^2_{H^{r+1}}]\}.\end{aligned}$$

Proof. Under the assumptions of the corollary, the operators $S_\omega, S_\omega^* : L^2(D) \to H^2(D) \cap H_0^1(D)$ are bounded. Using first the Aubin-Nitsche duality argument and then Céa's Lemma (see e.g. [32]), for the first term on the right hand side of (17), we find

$$\begin{aligned}\mathbb{E}[\|p(u^*) - \widetilde{p}^h(u^*)\|^2] &\le C\mathbb{E}[h^2\|p(u^*) - \widetilde{p}^h(u^*)\|^2_{H_0^1}] \\ &\le C\mathbb{E}[h^{2+2r}|p(u^*)|^2_{H^{r+1}}].\end{aligned}$$

A similar argument holds for the second term on the right hand side of (17). Finally the third term on the right hand side of (17) can be bounded directly by

$$h^2\mathbb{E}[\|y(u^*) - y^h(u^*)\|^2_{H_0^1}] \le Ch^2\mathbb{E}[Ch^{2r}|y|^2_{H^{r+1}}].$$

All these inequalities added together lead to the claim. □

4. Approximation in probability space

In this section we consider the semi-discrete (approximation in probability only) optimal control problem obtained by replacing the exact expectation $\mathbb{E}[\cdot]$ in (3) by a suitable quadrature formula $\widehat{E}[\cdot]$. The semi-discrete collocation problem then reads:

$$\begin{cases} \min_{u\in U} \widehat{J}(u) = \frac{1}{2}\widehat{E}[\|y_\omega(u) - z_d\|^2] + \frac{\alpha}{2}\|u\|^2 \\ \text{s.t.} \quad y_{\omega_i}(u) \in Y \quad \text{and} \\ b_{\omega_i}(y_{\omega_i}(u), v) = \langle g+u, v\rangle \quad \forall v \in Y \quad i = 1,\dots,N. \end{cases} \tag{19}$$

This quadrature formula could either be based on deterministic quadrature points or randomly distributed points leading, in this case, to a Monte Carlo type approximation. In particular, if $X : \Gamma \to \mathbb{R}$, $\omega \mapsto X(\omega)$, is a random variable, let $\widehat{E}[X] = \sum_{i=1}^N \zeta_i X(\omega_i)$ be the quadrature operator, where ζ_i are the quadrature weights and ω_i the quadrature knots. In the case of a Monte Carlo approximation, we have $\zeta_i = \frac{1}{N}$ for every i, and ω_i being independent and identically distributed (iid) points in Γ, all distributed according to the measure P.

In the next sub-sections we will particularize results for the cases of a Monte Carlo type quadrature. Although for the sake of notation we present these results for the semi-discrete problem (i.e. continuous in space, discrete in probability), they extend straightforwardly to the fully discrete problem in probability and in space, using a control $\widehat{u}^h$ instead of $\widehat{u}$, a solution of (19). We study the deterministic Gaussian-type quadrature method in the appendix.

4.1. **Monte Carlo method.** Consider a Monte Carlo approximation of the expectation appearing in (10), namely the exact expectation $\mathbb{E}$ is replaced by $E_{MC}^{\overrightarrow{\omega}}[X(\omega)] := \frac{1}{N}\sum_{i=1}^N X(\omega_i)$, where N denotes the number of ω_i, $i = 1,\dots,N$, of the random variable ω and denote by $\overrightarrow{\omega} = \{\omega_i\}_{i=1}^N$ the collection of these ω_i. We recall that the use of MC type approximations might be advantageous over a collocation/quadrature approach in cases where p is rough, which is, for example, the case when $a(\cdot,\cdot)$ is a rough random field w.r.t. the random parameter ω or has a short correlation length.

Remark 4. *We stress here that $\widehat{u}$ is a stochastic function because it depends on the N iid realizations $\overrightarrow{\omega} = \{\omega_i\}_{i=1}^N$ of the random variable ω.*

Theorem 4. *Let $\widehat{u}^*$ be the optimal control of problem (19) with $\widehat{E} = E_{MC}^{\overrightarrow{\omega}}$ and u^* be the exact optimal control of the continuous problem (10), then we have*

$$\frac{\alpha}{2}\mathbb{E}[\|\widehat{u}^* - u^*\|^2] + \mathbb{E}[\|y(u^*) - y(\widehat{u}^*)\|^2] \leq \frac{1}{N}\frac{1}{2\alpha}\mathbb{E}[\|p(\widehat{u}^*)\|^2].$$

Proof. Similarly to the proof of Theorem 2, the two optimality conditions read

$$\langle \nabla J(u^*), v_1 - u^*\rangle \geq 0 \quad \forall v_1 \in U \tag{20}$$

and

$$\langle \nabla J_{MC}(\widehat{u}^*), v_2 - \widehat{u}^*\rangle \geq 0 \quad \forall v_2 \in U \tag{21}$$

with

$$\nabla J_{MC}(\widehat{u}^*) = \alpha\widehat{u}^* + E_{MC}^{\overrightarrow{\omega}}[p(\widehat{u}^*)] \quad p(\widehat{u}^*) := S^*(S\widehat{u}^* - z).$$

Choosing $v_1 = \widehat{u}^*$ in (20) and $v_2 = u^*$ in (21) we obtain:

$$\langle \alpha(u^* - \widehat{u}^*) + \mathbb{E}[p(u^*)] - E_{MC}^{\overrightarrow{\omega}}[p(\widehat{u}^*)], \widehat{u}^* - u^*\rangle \geq 0,$$

which implies

$$\alpha\|u^* - \widehat{u}^*\|^2 \leq \langle \mathbb{E}[p(u^*)] - E_{MC}^{\overrightarrow{\omega}}[p(u^*)] + E_{MC}^{\overrightarrow{\omega}}[p(u^*)] - E_{MC}^{\overrightarrow{\omega}}[p(\widehat{u}^*)], \widehat{u}^* - u^*\rangle. \tag{22}$$

We can split the right hand side of (22) into two parts:

$$\langle \mathbb{E}[p(u^*)] - E_{MC}^{\overrightarrow{\omega}}[p(u^*)], \widehat{u}^* - u^* \rangle \leq \frac{1}{2\alpha}\|\mathbb{E}[p(u^*)] - E_{MC}^{\overrightarrow{\omega}}[p(u^*)]\|^2 + \frac{\alpha}{2}\|\widehat{u}^* - u^*\|^2$$

Moreover, for every $i = 1, \cdots, N$

$$\begin{aligned}\langle \widehat{u}^* - u^*, p_{\omega_i}(u^*) - p_{\omega_i}(\widehat{u}^*)\rangle &= b_{\omega_i}(y_{\omega_i}(\widehat{u}^*) - y_{\omega_i}(u^*), p_{\omega_i}(u^*) - p_{\omega_i}(\widehat{u}^*)) \\ &= \langle y_{\omega_i}(u^*) - y_{\omega_i}(\widehat{u}^*), y_{\omega_i}(\widehat{u}^*) - y_{\omega_i}(u^*)\rangle \\ &= -\|y_{\omega_i}(u^*) - y_{\omega_i}(\widehat{u}^*)\|^2,\end{aligned}$$

leading to

$$\langle \widehat{u}^* - u^*, E_{MC}^{\overrightarrow{\omega}}[p(u^*)] - E_{MC}^{\overrightarrow{\omega}}[p(\widehat{u}^*)]\rangle \leq -E_{MC}^{\overrightarrow{\omega}}[\|y(u^*) - y(\widehat{u}^*)\|^2]$$

We finally take the expectation of (22), w.r.t. the random sample $\overrightarrow{\omega} = \{\omega_i\}_{i=1}^N$ and exploit the fact that the Monte Carlo estimator is unbiased, that is $\mathbb{E}[E_{MC}^{\overrightarrow{\omega}}[X(\omega)]] = \mathbb{E}[X]$ for a random variable $X : \Gamma \to \mathbb{R}$.

$$\begin{aligned}\mathbb{E}[\frac{\alpha}{2}\|\widehat{u}^* - u^*\|^2 + E_{MC}^{\overrightarrow{\omega}}[\|y(u^*) - y(\widehat{u}^*)\|^2] &= \frac{\alpha}{2}\mathbb{E}[\|\widehat{u}^* - u^*\|^2 + \mathbb{E}[\|y(u^*) - y(\widehat{u}^*)\|^2] \\ &\leq \frac{1}{2\alpha}\mathbb{E}[\|\mathbb{E}[p(\widehat{u}^*)] - E_{MC}^{\overrightarrow{\omega}}[p(\widehat{u}^*)]\|^2] \\ &\leq \frac{1}{2\alpha}\mathbb{E}[\|\frac{1}{N}\sum_{i=1}^N p_{\omega_i}(\widehat{u}^*) - \mathbb{E}[p(\widehat{u}^*)]\|^2] \\ &\leq \frac{1}{2\alpha}\mathbb{E}[\frac{1}{N^2}\sum_{i=1}^N \|p_{\omega_i}(\widehat{u}^*) - \mathbb{E}[p(\widehat{u}^*)]\|^2] \\ &\leq \frac{1}{2\alpha}\frac{1}{N}\mathbb{E}[\|p(\widehat{u}^*) - \mathbb{E}[p(\widehat{u}^*)]\|^2] \\ &\leq \frac{1}{2\alpha}\frac{1}{N}\mathbb{E}[\|p(\widehat{u}^*)\|^2]\end{aligned}$$

what finishes the proof of the theorem. □

Theorem 4 shows that the semi-discrete optimal control $\widehat{u}^*$ converges at the usual MC rate of $1/\sqrt{N}$ in the root mean squared sense, with the constant being proportional to $\sqrt{\mathbb{E}[\|p(\widehat{u}^*)\|^2]}$.

5. Steepest descent method for fully discrete problem

Now we focus on a class of optimization methods to approximate the fully discrete minimization problem, using the Monte Carlo estimator to approximate the expectation in (11)

$$(23) \qquad \begin{cases} \min_{u^h \in U^h} J_{MC}(u^h) = \frac{1}{2}E_{MC}^{\overrightarrow{\omega}}[\|y_\omega^h(u^h) - z_d\|^2] + \frac{\alpha}{2}\|u^h\|^2 \\ \text{s.t.} \quad y_\omega^h(u^h) \in Y^h \quad \text{and} \\ b_\omega(y_\omega^h(u^h), v^h) = \langle g + u^h, v^h\rangle \quad \forall v^h \in Y^h, \quad \text{for a.e. } \omega \in \Gamma. \end{cases}$$

Specifically, we consider a simple gradient methods. The gradient method reads:

$$(24) \qquad \widehat{u}_{j+1}^h = \widehat{u}_j^h - \tau E_{MC}^{\overrightarrow{\omega}}[\nabla f^h(\widehat{u}_j^h, \omega)],$$

where $f^h(u,\omega) = \frac{1}{2}\|y_\omega^h(u) - z_d\|^2 + \frac{\alpha}{2}\|u^h\|^2$. Here, the index j represents the j-th iteration in the optimization recursion (24), while the superscript h denotes that we discretize the control u as well as the underlying PDE using Finite Elements on a fixed mesh of characteristic size h.

We first analyze the convergence of the continuous version of (24), i.e. of

$$u_{j+1} = u_j - \tau \mathbb{E}[\nabla f(u_j, \omega)] \ . \tag{25}$$

For this we prove a Lipschitz and a strong convexity condition for the function $f(u, \omega)$ for a.e. $\omega \in \Gamma$; which is still valid when replacing $f(u, \omega)$ by its discrete version $f^h(u^h, \omega_i)$.

Lemma 4 (Lipschitz condition)**.** *For the elliptic problem* (4) *and* $f(u, \omega)$ *as in* (3) *it holds that:*

$$\|\nabla f(u_1, \omega) - \nabla f(u_2, \omega)\| \leq L\|u_1 - u_2\| \quad \forall u_1, u_2 \in U \text{ and a.e. } \omega \in \Gamma, \tag{26}$$

with $L = \alpha + \frac{C_p^4}{a_{min}^2}$*, where* C_p *is the Poincaré constant. For the Finite Element approximation as in* (11) *the same inequality holds with the same constant*

$$\|\nabla f^h(u_1^h, \omega) - \nabla f^h(u_2^h, \omega)\| \leq L\|u_1^h - u_2^h\| \quad \forall u_1^h, u_2^h \in U^h \text{ and a.e. } \omega \in \Gamma.$$

Proof. For a.e. $\omega \in \Gamma$, and every $u, u' \in U$ we have that

$$\nabla f(u', \omega) - \nabla f(u, \omega) = \alpha(u' - u) + p_\omega(u') - p_\omega(u), \tag{27}$$

and

$$\begin{aligned}\|p_\omega(u') - p_\omega(u)\|^2 &\leq C_p^2 \|\nabla_x p_\omega(u') - \nabla_x p_\omega(u)\|^2 \\ &\leq \frac{C_p^2}{a_{min}} b_\omega\big(p_\omega(u') - p_\omega(u), p_\omega(u') - p_\omega(u)\big) \\ &\leq \frac{C_p^2}{a_{min}} \langle p_\omega(u') - p_\omega(u), y_\omega(u') - y_\omega(u)\rangle \\ &\leq \frac{C_p^2}{a_{min}} \|p_\omega(u') - p_\omega(u)\| \|y_\omega(u') - y_\omega(u)\|.\end{aligned}$$

With same arguments we find that

$$\begin{aligned}\|y_\omega(u') - y_\omega(u)\|^2 &\leq C_p^2 \|\nabla_x y_\omega(u') - \nabla_x y_\omega(u)\|^2 \\ &\leq \frac{C_p^2}{a_{min}} b_\omega\big(y_\omega(u') - y_\omega(u), y_\omega(u') - y_\omega(u)\big) \\ &\leq \frac{C_p^2}{a_{min}} \langle y_\omega(u') - y_\omega(u), u' - u\rangle \\ &\leq \frac{C_p^2}{a_{min}} \|y_\omega(u') - y_\omega(u)\| \|u' - u\|.\end{aligned}$$

Combining (27) with the two last estimates, we find

$$\begin{aligned}\|\nabla f(u', \omega) - \nabla f(u, \omega)\| &\leq \alpha\|u' - u\| + \|p_\omega(u') - p_\omega(u)\| \\ &\leq \Big(\alpha + \frac{C_p^4}{a_{min}^2}\Big)\|u' - u\|.\end{aligned}$$

The proof in the FE setting follows verbatim the above one. □

Lemma 5 (Strong Convexity)**.** *For the elliptic problem* (4) *and* $f(u, \omega)$ *as in* (3) *it holds that*

$$\frac{l}{2}\|u_1 - u_2\|^2 \leq \langle \nabla f(u_1, \omega) - \nabla f(u_2, \omega), u_1 - u_2\rangle \quad \forall u_1, u_2 \in U \text{ and a.e. } \omega \in \Gamma, \tag{28}$$

with $l = 2\alpha$*. The same estimate holds for the FE approximation as in* (11)*, namely:*

$$\frac{l}{2}\|u_1^h - u_2^h\|^2 \leq \langle \nabla f^h(u_1^h, \omega) - \nabla f^h(u_2^h, \omega), u_1^h - u_2^h\rangle \quad \forall u_1^h, u_2^h \in U^h \text{ and a.e. } \omega \in \Gamma.$$

Proof. For every $\omega \in \Gamma$, and every $u, u' \in U$:

$$\begin{aligned}\langle u'-u, \nabla f(u',\omega) - \nabla f(u,\omega)\rangle &= \langle u'-u, \alpha(u'-u) + p_\omega(u') - p_\omega(u)\rangle \\ &= \alpha\|u'-u\|^2 + \langle u'-u, p_\omega(u') - p_\omega(u)\rangle \\ &= \alpha\|u'-u\|^2 + b_\omega\big(y_\omega(u') - y_\omega(u), p_\omega(u') - p_\omega(u)\Big) \\ &= \alpha\|u'-u\|^2 + \langle y_\omega(u') - y_\omega(u), y_\omega(u') - y_\omega(u)\rangle \\ &= \alpha\|u'-u\|^2 + \|y_\omega(u') - y_\omega(u)\|^2 \\ &\geq \alpha\|u'-u\|^2\end{aligned}$$

The same proof applies to the FE case. □

Based on the results of Lemmas 4 and 5, it is straightforward to show the convergence of the iterates. We state the result for the gradient method for the continuous problem (25) in the following Lemma and the result for the fully discretized problem(24) in Theorem 5.

Lemma 6. *Let u^* be the optimal solution of the control problem* (10) *and $\{u_j\}_{j\in\mathbb{N}}$ the iterations produced by* (25)*. Then for any $0 < \tau < l/L^2$ we have*

$$\|u_{j+1} - u^*\|^2 \leq (1 - \tau l + \tau^2 L^2)\|u_j - u^*\|^2 \leq (1 - \tau l + \tau^2 L^2)^{j+1}\|u_0 - u^*\|^2, \tag{29}$$

and $\|u_j - u^*\| \to 0$ as $j \to \infty$.

Proof. Since u^* satisfies the optimality condition $\nabla J(u^*) = 0$ we have

$$u_{j+1} - u^* = u_j - u^* - \tau\mathbb{E}[\nabla f(u_j,\omega) - \nabla f(u^*,\omega)].$$

Consequently,

$$\begin{aligned}\|u_{j+1} - u^*\|^2 &= \|u_j - u^*\|^2 + \tau^2\|\mathbb{E}[\nabla f(u_j,\omega) - \nabla f(u^*,\omega)]\|^2 \\ &\quad - 2\tau\langle u_j - u^*, \mathbb{E}[\nabla f(u_j,\omega) - \nabla f(u^*,\omega)]\rangle \\ &\leq (1 - \tau l + \tau^2 L^2)\|u_j - u^*\|^2.\end{aligned}$$

The condition $0 < \tau < l/L^2$ guarantees that $0 < 1 - \tau l + \tau^2 L^2 < 1$ and the claim follows. □

As mentioned before, we now provide an error bound for the approximate solution $\widehat{u}_j^h$ defined in (24), as a function of the discretization parameters j, h, and N.

Theorem 5. *Let $\widehat{u}_j^h$ be the solution produced by* (24) *at the j-th iteration and denote by u^* the solution of the optimal problem* (10)*. Then under the assumptions of Corollary 1, there exist constants $C_1, C_2, C_3 > 0$ such that*

$$\mathbb{E}[\|\widehat{u}_j^h - u\|^2] \leq C_1 e^{-\rho j} + \frac{C_2}{N} + C_3 h^{2r+2}\ , \tag{30}$$

with $\rho = -\log(1 - \tau l + \tau^2 L^2)$ for $0 < \tau < l/L^2$.

Proof. The global error can be decomposed as follows:

$$\mathbb{E}[\|\widehat{u}_j^h - u^*\|^2] \leq 3\underbrace{\mathbb{E}[\|\widehat{u}_j^h - \widehat{u}^{h,*}\|^2]}_{\text{gradient}} + 3\underbrace{\mathbb{E}[\|\widehat{u}^{h,*} - u^{h*}\|^2]}_{\text{MC}} + 3\underbrace{\mathbb{E}[\|u^{h*} - u^*\|^2]}_{\text{FE error}}.$$

The first term $\mathbb{E}[\|\widehat{u}_j^h - \widehat{u}^{h,*}\|^2]$ quantifies the convergence of the finite dimensional steepest descent algorithm and can be estimated as in Lemma 6. In fact, for any sample $\overrightarrow{\omega} = \{\omega_i\}_{i=1}^N$ we have

$$\|\widehat{u}_j^h - \widehat{u}^{h,*}\|^2 \leq (1 - \tau l + \tau^2 L^2)^j\|\widehat{u}_0^h - \widehat{u}^{h,*}\|^2 = e^{-\rho j}\|\widehat{u}_0^h - \widehat{u}^{h,*}\|^2.$$

with $\rho = -\log(1 - \tau l + \tau^2 L^2)$. Hence taking expectation w.r.t. $\overrightarrow{\omega}$,

$$\mathbb{E}[\|\widehat{u}_j^h - \widehat{u}^{h,*}\|^2] \leq e^{-\rho j}\mathbb{E}[\|\widehat{u}_0^h - \widehat{u}^{h,*}\|^2].$$

The second term $\mathbb{E}[\|\widehat{u}^{h,*} - u^{h*}\|^2]$ accounts for the standard MC error and can be controlled as in Theorem 4 (applied on the FE approximation) leading to

$$\mathbb{E}[\|\widehat{u}^{h,*} - u^{h*}\|^2] \leq \frac{1}{\alpha^2 N}\mathbb{E}[\|p(\widehat{u}^h)\|^2].$$

Finally, the term $\mathbb{E}[\|u^{h*} - u^*\|^2]$ can be controlled by the result in Corollary 1, namely by

$$\|u^{h*} - u^*\|^2 \leq C\big(\mathbb{E}[|y_\omega(u^*)|^2_{H^{r+1}}] + \mathbb{E}[|p_\omega(u^*)|^2_{H^{r+1}}]\big)h^{2r+2},$$

so that the claim follows. □

We conclude this Section by analyzing the complexity of the Algorithm 1 based on the optimization scheme (24). We assume that the primal and dual problems can be solved, using a triangulation with mesh size h, in computational time $C_h = O(h^{-n\gamma})$. Here, $\gamma \in [1, 3]$ is a parameter representing the efficiency of the linear solver used (e.g. $\gamma = 3$ for a direct solver and $\gamma = 1$ up to a logarithm factor for an optimal multigrid solver), while n is the dimension of the physical space. Hence the overall computational work W of j gradient iterations is proportional to $W \simeq 2Njh^{-n\gamma}$.

Corollary 2. *In order to achieve a given tolerance $O(tol)$, i.e. to guarantee that $\mathbb{E}[\|\widehat{u}_j^h - u\|^2] \lesssim tol^2$, the total required computational work is bounded by*

$$W \lesssim tol^{-2-\frac{n\gamma}{r+1}}|\log(tol)|.$$

Proof. To achieve a tolerance $O(tol)$, we can equidistribute the precision tol^2 over the three terms in (30). This leads to the choices given in Algorithm 1:

$$j_{max} \simeq -\log(tol), \quad h \simeq tol^{\frac{1}{r+1}}, \quad N \simeq tol^{-2}.$$

Hence the total cost for computing a solution $\widehat{u}_{j_{max}}^h$ that achieves the required tolerance is $W \simeq 2Nj_{max}h^{-n\gamma} = tol^{-2-\frac{n\gamma}{r+1}}|\log(tol)|$ as claimed. □

We propose a description of the algorithm used in this Section, in Algorithm 1.

Algorithm 1: Steepest descent method for fully discrete problem

```
Data:
Given a desired tolerance $tol$:
Choose $\tau < \frac{l}{L^2}$, $j_{max} \simeq -\log(tol)$, $N_{MC} \simeq tol^{-2}$, $h \simeq tol^{\frac{1}{r+1}}$
Generate $N_{MC}$ iid realizations of the random field $a_i = a(\cdot, \omega_i)$,
$i = 1, \dots, N_{MC}$.
initialization:
$u = 0$;
for $j = 1, \dots, j_{max}$ do
    $\widehat{p} = 0$;
    for $i = 1, \dots, N_{MC}$ do
        solve primal problem by FE $\rightarrow y(a_i, u)$
        solve dual problem by FE $\rightarrow p(a_i, u)$
        update $\widehat{p} = \widehat{p} + p(a_i, u)/N_{MC}$
    end
    $\widehat{\nabla J} = \alpha u + \widehat{p}$
    $u = u - \tau\widehat{\nabla J}$
end
```

The second (MC) term in the error bound (30) C_2/N is numerically a problem/limitation to compute efficiently a solution. That is why in the following Section we combine the first two terms, using Stochastic Gradient techniques.

6. Stochastic Gradient with fixed mesh size.

As an alternative to the fixed MC gradient method (24) considered in Section 5, in which the sample size N is fixed beforehand and a full sample average is computed at each iteration, here we consider a variant, known in literature as Stochastic Approximation (SA) or Stochastic Gradient (SG) [14, 31, 33, 38, 39].

The classic version of such a method, the so-called Robbins-Monro method, works as follows. Within the steepest descent algorithm the exact gradient $\nabla J = \nabla\mathbb{E}[f] = \mathbb{E}[\nabla f]$ is replaced by $\nabla f(\cdot, \omega_j)$, where the random variable ω_j is re-sampled independently at each iteration of the steepest-descent method:

$$u_{j+1} = u_j - \tau_j \nabla f(u_j, \omega_j). \tag{31}$$

Here, τ_j is the step-size of the algorithm and is decreasing as $1/j$ in the usual approach. We consider a generalization of this method, in which the pointwise gradient $\nabla f(\cdot, \omega_j)$ is replaced by a sample average over N_j iid realizations which are drawn independently of the previous iterations. More precisely, let $\overrightarrow{\omega_j} = (\omega_j^{(1)}, \cdots, \omega_j^{(N_j)})$, then we define the recursion as

$$u_{j+1} = u_j - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \omega)], \tag{32}$$

where $E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u,\omega)] = \frac{1}{N_j}\sum_{i=1}^{N_j} \nabla f(u, \omega_j^{(i)})$ is the usual Monte Carlo estimator. Notice that the Robbins-Monro method is a special case of this scheme, namely with $N_j = 1$ for all j. In what follows, we investigate optimal choices of the sequences $\{\tau_j\}_j$ and $\{N_j\}_j$, and the overall computational complexity of the corresponding algorithm. First we analyze the continuous version (i.e. no Finite Element discretization).

Theorem 6. *Let u^* be the solution of the continuous OCP* (10) *and denote by u_j the j-th iterate of* (32). *Then it holds that*

$$\mathbb{E}[\|u_{j+1} - u^*\|^2] \leq c_j \mathbb{E}[\|u_j - u^*\|^2] + \frac{2\tau_j^2}{N_j}\mathbb{E}[\|\nabla f(u^*, \omega)\|^2], \tag{33}$$

with $c_j := 1 - \tau_j l + L^2\Big(1 + \frac{2}{N_j}\Big)\tau_j^2$.

Proof. Using inequalities (26) and (28), we can formulate a recursive formula to control the error between successive iterations. As each iteration uses an independent sample, we need to keep track of the history of the sampling $\omega_{[j-1]} = \{\overrightarrow{\omega_1}, \ldots, \overrightarrow{\omega_{j-1}}\}$ to be able to define u_j. Thus we introduce the conditional expectation $G[\cdot] = \mathbb{E}[\cdot|\omega_{[j-1]}]$. Using $\mathbb{E}[\nabla f(u^*, \omega)] = 0$, we have:

$$\begin{aligned} u_{j+1} - u^* &= u_j - u^* - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \overrightarrow{\omega_j})] + \tau_j \mathbb{E}[\nabla f(u^*, \omega)] \\ &= u_j - u^* - \tau_j G[\nabla f(u_j, \omega)] + \tau_j \mathbb{E}[\nabla f(u^*, \omega)] + \tau_j \big(G[\nabla f(u_j, \omega)] - E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \overrightarrow{\omega_j})]\big) \\ &= u_j - u^* - \tau_j T_1 + \tau_j T_2, \end{aligned}$$

with $T_1 := G[\nabla f(u_j, \omega)] - \mathbb{E}[\nabla f(u^*, \omega)]$ and $T_2 := G[\nabla f(u_j, \omega)] - E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \overrightarrow{\omega_j})]$. Hence,

$$\begin{aligned} \|u_{j+1} - u^*\|^2 =& \|u_j - u^*\|^2 + \tau_j^2\|T_1\|^2 + \tau_j^2\|T_2\|^2 \\ & - 2\tau_j\langle u_j - u^*, T_1\rangle + 2\tau_j\langle u_j - u^*, T_2\rangle - 2\tau_j^2\langle T_1, T_2\rangle. \end{aligned}$$

Moreover, by definition of T_1, we find:

$$\begin{aligned}
\|T_1\|^2 &= \|G[\nabla f(u_j,\omega)] - \mathbb{E}[\nabla f(u^*,\omega)]\|^2 \\
&= \|G[\nabla f(u_j,\omega) - \nabla f(u^*,\omega)]\|^2 \qquad [\overrightarrow{\omega_j} \text{ being independent of } \omega_{[j-1]}] \\
&= \int_D \big(G[\nabla f(u_j,\omega) - \nabla f(u^*,\omega)]\big)^2 \mathrm{d}x \\
&\leq \int_D G[|\nabla f(u_j,\omega) - \nabla f(u^*,\omega)|^2]\mathrm{d}x \qquad [\text{Jensen}'\text{s inequality}] \\
&= G[\|\nabla f(u_j,\omega) - \nabla f(u^*,\omega)\|^2] \\
&\leq L^2 G[\|u_j - u^*\|^2],
\end{aligned}$$

where we have used Jensen's inequality for conditional expectation: $\phi(G[X]) \leq G[\phi(X)]$ for ϕ convex. See e.g.[42].
Then taking the expectation over all the history sampling $\omega_{[j-1]}$, we have:

$$\begin{aligned}
\mathbb{E}[\|T_1\|^2] &\leq L^2\mathbb{E}[G[\|u_j - u^*\|^2]] \\
&= L^2\mathbb{E}[\|u_j - u^*\|^2],
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[\langle u_j - u^*, T_1\rangle] &= \mathbb{E}[\langle u_j - u^*, G[\nabla f(u_j,\omega) - \nabla f(u^*,\omega)]\rangle] \\
&= \mathbb{E}[G[\langle u_j - u^*, \nabla f(u_j,\omega) - \nabla f(u^*,\omega)\rangle]] \\
&\geq \mathbb{E}[G[\frac{l}{2}\|u_j - u^*\|^2]] \qquad\qquad [\text{Strong Convexity (28)}] \\
&= \frac{l}{2}\mathbb{E}[\|u_j - u^*\|^2].
\end{aligned}$$

Concerning the term T_2, it holds that,

$$\|T_2\|^2 = \|G\left[\nabla f(u_j,\omega)\right] - E_{MC}^{\overrightarrow{\omega_j}}\left[\nabla f(u_j,\omega)\right]\|^2 = \int_D \Big(G[\nabla f(u_j,\omega)] - E_{MC}^{\overrightarrow{\omega_j}}\left[\nabla f(u_j,\omega)\right]\Big)^2 \mathrm{d}x.$$

Again, taking the expectation w.r.t. $\omega_{[j]}$ yields

$$\begin{aligned}
\mathbb{E}\left[\|T_2\|^2\right] &= \mathbb{E}\left[\int_D \left(\frac{1}{N_j}\sum_{i=1}^{N_j}\Big(G[\nabla f(u_j,\omega)] - \nabla f\left(u_j,\omega_j^{(i)}\right)\Big)\right)^2 \mathrm{d}x\right] \\
&= \mathbb{E}\left[\int_D \frac{1}{N_j^2}\sum_{i,l=1}^{N_j}\Big(\nabla f\left(u_j,\omega_j^{(i)}\right) - G\left[\nabla f\left(u_j,\omega\right)\right]\Big)\Big(\nabla f\left(u_j,\omega_j^{(l)}\right) - G\left[\nabla f\left(u_j,\omega\right)\right]\Big)\mathrm{d}x\right] \\
&= \int_D \frac{1}{N_j^2}\sum_{i,l=1}^{N_j}\mathbb{E}\left[\Big(\nabla f\left(u_j,\omega_j^{(i)}\right) - G\left[\nabla f\left(u_j,\omega\right)\right]\Big)\Big(\nabla f\left(u_j,\omega_j^{(l)}\right) - G\left[\nabla f\left(u_j,\omega\right)\right]\Big)\right]\mathrm{d}x \\
&= \int_D \frac{1}{N_j^2}\sum_{i,l=1}^{N_j}\mathbb{E}\left[G\left[\Big(\nabla f\left(u_j,\omega_j^{(i)}\right) - G\left[\nabla f(u_j,\omega)\right]\Big)\Big(\nabla f\left(u_j,\omega_j^{(l)}\right) - G[\nabla f(u_j,\omega)]\Big)\right]\right]\mathrm{d}x.
\end{aligned}$$

Observe that, conditional upon $\omega_{[j-1]}$, the random variables $Y_i = \nabla f(u_j,\omega_j^{(i)}) - G[\nabla f(u_j,\omega)]$, $i = 1,\ldots,N_j$, are mutually independent and have zero mean, i.e.

$\mathbb{E}\left[Y_i|\omega_{[j-1]}\right] = G[Y_i] = 0$ and $G(Y_iY_j) = 0$ when $i \neq j$. Therefore it follows that

$$\begin{aligned}
\mathbb{E}\left[\|T_2\|^2\right] &= \int_D \frac{1}{N_j^2}\sum_{i=1}^{N_j}\mathbb{E}\left[G\left[\left(\nabla f\left(u_j,\omega_j^{(i)}\right) - G\left[\nabla f\left(u_j,\omega\right)\right]\right)^2\right]\right]\mathrm{d}x \\
&= \mathbb{E}\left[\int_D \frac{1}{N_j}G\left[\left(\nabla f\left(u_j,\omega\right) - G\left[\nabla f(u_j,\omega)\right]\right)^2\right]\mathrm{d}x\right] \\
&\leq \mathbb{E}\left[\int_D \frac{1}{N_j}G\left[\nabla f^2(u_j,\omega)\right]\mathrm{d}x\right] \\
&= \frac{1}{N_j}\mathbb{E}\left[\|\nabla f(u_j,\omega)\|^2\right] \\
&\leq \frac{2}{N_j}\mathbb{E}\left[\|\nabla f(u_j,\omega) - \nabla f(u^*,\omega)\|^2 + \|\nabla f(u^*,\omega)\|^2\right] \qquad \text{[Lipschitz condition (26)]} \\
&\leq \frac{2L^2}{N_j}\mathbb{E}\left[\|u_j - u^*\|^2\right] + \frac{2}{N_j}\mathbb{E}\left[\|\nabla f(u^*,\omega)\|^2\right].
\end{aligned}$$

Finally, we have that

$$\begin{aligned}
\mathbb{E}[\langle u_j - u^*, T_2\rangle] &= \mathbb{E}[G[\langle u_j - u^*, T_2\rangle]] \\
&= \mathbb{E}[\langle u_j - u^*, G[T_2]\rangle] \\
&= \frac{1}{N_j}\sum_{i=1}^{N_j}\mathbb{E}[\langle u_j - u^*, G[Y_i]\rangle] \\
&= 0,
\end{aligned}$$

and, similarly, $\mathbb{E}[\langle T_1, T_2\rangle] = \mathbb{E}[G[\langle T_1, T_2\rangle]] = \mathbb{E}[\langle T_1, G[T_2]\rangle] = 0$, which concludes the proof. □

We now consider the FE version of (32) and focus on the common setting $(\tau_j, N_j) = (\tau_0/j, \overline{N})$, which is a generalization of Robbins-Monro method:

$$u_{j+1}^h = u_j^h - \frac{\tau_0}{j}E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^h(u_j^h,\omega)] \tag{34}$$

with $\overrightarrow{\omega_j} := (\omega_j^{(1)}, \cdots, \omega_j^{(\overline{N})})$.

Theorem 7. *Suppose that the assumptions of Corollary 1 hold and let u_j^h denote the j-th iterate of* (34). *For the choice $(\tau_j, N_j) = (\tau_0/j, \overline{N})$ with $\tau_0 > 1/l$ we have*

$$\mathbb{E}[\|u_j^h - u^*\|^2] \leq D_1 j^{-1} + D_2 h^{2r+2}\ , \tag{35}$$

for suitable constants $D_1, D_2 > 0$ independent of j and h.

Proof. The factor c_j in (33) becomes in this case

$$c_j = 1 - \frac{\tau_0 l}{j} + \frac{\tau_0^2 L^2}{j^2}\Big(1 + \frac{2}{\overline{N}}\Big).$$

We use the recursive formula (33) and set as before u^{h*} the exact optimal control for the FE problem defined in (11). We emphasize that (11) has no approximation in the probability space. Setting $a_j = \mathbb{E}[\|u_j^h - u^{h*}\|^2]$ and $\beta_j = \frac{2\tau_j^2}{N_j}\mathbb{E}[\|\nabla f(u^{h*},\omega)\|^2]$, from (33) applied to the sequence of Finite Element solutions $\{u_j^h\}_{j>0}$ we find

$$
\begin{aligned}
a_{j+1} \leq & c_j a_j + \beta_j \\
\leq & c_j c_{j-1} a_{j-1} + c_j \beta_{j-1} + \beta_j \\
\leq & \cdots \\
\leq & \underbrace{\Big(\prod_{i=1}^{j} c_i\Big)}_{=\kappa_j} a_1 + \underbrace{\sum_{i=1}^{j} \beta_i \prod_{l=i+1}^{j} c_l}_{=\mathcal{B}_j} . \qquad (36)
\end{aligned}
$$

For the first term κ_j, computing its logarithm, we have,

$$
\log(\kappa_j) = \sum_{i=1}^{j} \log(1 - \frac{\tau_0 l}{i} + \frac{M}{i^2}) \leq \sum_{i=1}^{j} \frac{-\tau_0 l}{i} + \sum_{i=1}^{j} \frac{M}{i^2},
$$

where we have set $M = \tau_0^2 L^2 \big(1 + \frac{2}{N}\big)$. Thus

$$
\log(\kappa_j) \leq -\tau_0 l \log j + M', \ \text{ with } \ M' = \sum_{i=1}^{\infty} \frac{M}{i^2},
$$

and $\kappa_j \lesssim j^{-\tau_0 l}$. For the second term $\mathcal{B}_j$ in (36) we have:

$$
\mathcal{B}_j = \sum_{i=1}^{j} \beta_i \prod_{k=i+1}^{j} c_k \leq \sum_{i=1}^{j} \frac{S}{i^2} \underbrace{\prod_{k=i+1}^{j} \Big(1 - \frac{\tau_0 l}{k} + \frac{\tau_0^2 L^2}{k^2}\Big)}_{=K_{ij}}, \ \text{ with } \ S = \frac{2\tau_0^2}{N} \mathbb{E}[\|\nabla f(u^{h*}, \omega)\|^2].
$$

For the term K_{ij} we can proceed as follow:

$$
\begin{aligned}
\log(K_{ij}) &= \sum_{k=i+1}^{j} \log\left(1 - \frac{\tau_0 l}{k} + \frac{M}{k^2}\right) \\
&\leq \sum_{k=i+1}^{j} \Big(- \frac{\tau_0 l}{k} + \frac{M}{k^2}\Big) \\
&\leq -\tau_0 l (\log(j+1) - \log(i+1)) + M\left(\frac{1}{i} - \frac{1}{j}\right),
\end{aligned}
$$

which shows that

$$
K_{ij} \leq (j+1)^{-\tau_0 l} (i+1)^{\tau_0 l} \exp\left(M\left(\frac{1}{i} - \frac{1}{j}\right)\right).
$$

It follows that

$$
\begin{aligned}
\mathcal{B}_j &\leq (j+1)^{-\tau_0 l} \underbrace{\exp\left(-\frac{M}{j}\right)}_{\leq 1} \sum_{i=1}^{j} S i^{\tau_0 l - 2} \underbrace{\exp\left(\frac{M}{i}\right)}_{\leq \exp(M)} \\
&\leq S \exp(M) (j+1)^{-\tau_0 l} \sum_{i=1}^{j} i^{\tau_0 l - 2} \lesssim j^{-1},
\end{aligned}
$$

for $\tau_0 > 1/l$. Eventually, we obtain the following upper bound, for two constants $D_3 > 0$ and $D_4 > 0$:

$$
a_{j+1} \leq D_3 j^{-\tau_0 l} a_1 + D_4 j^{-1}. \qquad (37)
$$

From the condition $\tau_0 > \frac{1}{l}$, we conclude that

$$
a_{j+1} \leq D_1 j^{-1}, \qquad (38)
$$

with D_1 possibly depending in $\|u_0^h - u^{h*}\|$. Finally splitting the error as

$$\mathbb{E}[\|u_j^h - u^*\|^2] \leq 2\mathbb{E}[\|u_j^h - u^{h*}\|^2] + 2\mathbb{E}[\|u^{h*} - u^*\|^2],$$

and using (18) to bound the second term, the claim follows. □

We propose a description of the SG algorithm 2 with fixed mesh size, used in Section 6.

Algorithm 2: Stochastic Gradient with fixed mesh size algorithm, with $\overline{N} = 1$.

```
Data:
Given a desired tolerance $tol$, choose $\frac{1}{l} < \tau_0$, $j_{max} \simeq tol^{-2}$, and $h \simeq tol^{\frac{1}{r+1}}$
initialization:
$u = 0$;
for $j = 1, \dots, j_{max}$ do
    sample one realization $a_j = a(\cdot, \omega_j)$ of the random field
    solve primal problem → $y(a_j, u)$ using FE on mesh $h$
    solve dual problem → $p(a_j, u)$ using FE on mesh $h$
    $\widehat{\nabla J} = \alpha u + p(a_j, u)$
    $u = u - \tau_j \widehat{\nabla J}$
end
```

We conclude this section by analyzing the complexity of the Algorithm 2.

Corollary 3. *To achieve a given tolerance $O(tol)$, i.e. to guarantee that $\mathbb{E}[\|u_j^h - u^*\|^2] \lesssim tol^2$, the total required computational work is bounded by*

$$W \lesssim tol^{-2-\frac{n\gamma}{r+1}}.$$

Here, we recall that the primal and dual problems can be solved, using a triangulation with mesh size h, in computational time $C_h = O(h^{-n\gamma})$, and r is the degree of the continuous FE that we use.

Proof. To achieve a tolerance $O(tol^2)$ for the error $\mathbb{E}[\|u_j^h - u^*\|^2]$, we can equidistribute the precision tol^2 over the two terms in (35). This leads to the choice:

$$j_{max} \simeq tol^{-2}, \quad h \simeq tol^{\frac{1}{r+1}}.$$

The cost for solving one deterministic PDE with the FE method is proportional to $h^{-n\gamma}$. Hence the total cost for computing a solution u_j^h that achieves the required tolerance is

$$W \simeq 2\overline{N} j h^{-n\gamma} = O(tol^{-2-\frac{n\gamma}{r+1}}),$$

as claimed. □

Remark 5. *Other choices of (τ_j, N_j) have been investigated. For example we have studied the SG with step-size $\tau_j = \tau_0/j$, $\tau_0 l - 1 > 0$ and increasing the MC sample size $N_j \sim j^{\tau_0 l - 1}$. With this choice the estimate in* (35) *becomes*

$$a_{j+1} \leq D_4 j^{-\tau_0 l} \log(j) \;, \tag{39}$$

which leads to the choice $j_{max} \simeq tol^{-\frac{2}{\tau_0 l}} |\log(tol)|^{\frac{1}{\tau_0 l}}$ and a final complexity

$$W \simeq 2 \sum_{i=1}^{j} i^{\tau_0 l - 1} h^{-n\gamma} \simeq 2 j^{\tau_0 l} h^{-n\gamma} = O(tol^{-2-\frac{n\gamma}{r+1}} |\log(tol)|).$$

The proof of the bound (39) *is detailed in Appendix B for completeness.*

Fixed MC gradient	SG - Variable step-size	SG - Variable step-size and N_j
$\tau_j = \tau_0$	$\tau_j = \tau_0/j$	$\tau_j = \tau_0/j$
$N \simeq tol^{-2}$	$N_j = \overline{N}$	$N_j = j^{\tau_0 l-1}$
$h \simeq tol^{\frac{1}{r+1}}$	$h \simeq tol^{\frac{1}{r+1}}$	$h \simeq tol^{\frac{1}{r+1}}$
$j_{max} \simeq -\log(tol)$	$j_{max} \simeq tol^{-2}$	$j_{max} \simeq tol^{-\frac{2}{\tau_0 l}} \lvert\log(tol)\rvert^{\frac{1}{\tau_0 l}}$
$W \lesssim tol^{-2-\frac{n\gamma}{r+1}} \lvert\log(tol)\rvert$	$W \lesssim tol^{-2-\frac{n\gamma}{r+1}}$	$W \lesssim tol^{-2-\frac{n\gamma}{r+1}} \lvert\log(tol)\rvert$

Table 1. Complexity analysis overview for different optimization methods

Remark 6. *Since the constant l may be challenging to estimate in practice, it is often difficult to fulfill the condition $\tau_0 > 1/l$. To bypass this difficulty, one could consider the* Averaged Stochastic Gradient method *[38] instead, in which the step size $\tau_j = \tau_0/j^\eta$, $\eta \in (0,1)$ is chosen, with $N_j = \overline{N}$ and the averaged control $\frac{1}{j}\sum_{i=1}^{j} u_i$ is considered. The analysis of this alternative method is postponed to a future work.*

Table 1 summarizes the results obtained in both the fixed sample size and increasing sample size regimes. There, the total work (W) to achieve a given tolerance (tol) is presented. We see from the table that the two considered SG versions improve the complexity only by a logarithmic factor compared to the fixed gradient algorithm. The advantage we see in the SG version w.r.t. the fixed gradient, is that we do not have to fix in advance the sample size N and we can just monitor the convergence of the SG iteration until a prescribed tolerance is reached. However, in Algorithm 2, we do have to choose in advance the FE mesh size. It is therefore natural to look at a further variation of the SG algorithm in which the FE mesh is refined during the iterations until a prescribed tolerance is reached. This is detailed in the next Section.

7. Stochastic Gradient with variable mesh size

In this section, we refine the mesh used for our FE approximation, while running the optimization routine. The new mesh size h_j is now depending on the iteration j. Here we study only sequences of nested meshes of size $h_j = 2^{-\ell(j)}$ with $\ell : \mathbb{N} \to \mathbb{N}$ being an increasing function. The optimization procedure then reads:

$$u_{j+1}^{h_{j+1}} = u_j^{h_j} - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \omega)], \tag{40}$$

with $\overrightarrow{\omega_j} := (\omega_j^{(1)}, \cdots, \omega_j^{(N_j)})$. Notice that if non-nested meshes are used, a projection operator should be added in (40) to transfer information from one mesh to another. We first derive an error recurrence formula in the spirit of (33) for the particular recurrence (40) with a decreasing mesh-size h_j.

Theorem 8. *Denoting by $u_{j+1}^{h_{j+1}}$ the approximated control obtained using the recursive definition* (40)*, and u^* the exact control for the continuous optimal problem* (10)*, we have:*

$$\begin{aligned} &\mathbb{E}[\|u_{j+1}^{h_{j+1}} - u^*\|^2] \\ &\quad \le c_j \mathbb{E}[\|u_j^{h_j} - u^*\|^2] + \frac{4\tau_j^2}{N_j}\mathbb{E}[\|\nabla f(u^*, \omega)\|^2] + 4\tau_j\big(\tau_j(1 + \frac{2}{N_j}) + \frac{1}{l}\big)Ch_j^{2r+2}, \end{aligned} \tag{41}$$

with $c_j = 1 - \frac{\tau_j l}{2} + \tau_j^2 L^2\big(2 + \frac{2}{N_j}\big)$.

Proof. Subtracting the optimal continuous control u^* from both sides of the recurrence formula (40), we get

$$\begin{aligned} u_{j+1}^{h_{j+1}} - u^* =& u_j^{h_j} - u^* - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \omega)] \pm \tau_j \mathbb{E}[\nabla f^{h_j}(u^*)] \pm \tau_j G[\nabla f^{h_j}(u_j^{h_j})] + \tau_j \mathbb{E}[\nabla f(u^*)] \\ =& u_j^{h_j} - u^* + \tau_j \Big(\mathbb{E}[\nabla f^{h_j}(u^*)] - G[\nabla f^{h_j}(u_j^{h_j})]\Big) \\ &+ \tau_j \Big(G[\nabla f^{h_j}(u_j^{h_j})] - E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \omega)]\Big) + \tau_j \left(\mathbb{E}[\nabla f(u^*) - \nabla f^{h_j}(u^*)]\right). \end{aligned}$$

Then setting as in proof of Theorem 6:

$$\begin{aligned} T_1 &:= G[\nabla f^{h_j}(u_j^{h_j})] - \mathbb{E}[\nabla f^{h_j}(u^*)], \\ T_2 &:= G[\nabla f^{h_j}(u_j^{h_j})] - E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \omega)], \\ T_3 &:= \mathbb{E}[\nabla f(u^*) - \nabla f^{h_j}(u^*)], \end{aligned}$$

we can rewrite the last equality as:

$$u_{j+1}^{h_{j+1}} - u^* = u_j^{h_j} - u^* - \tau_j T_1 + \tau_j T_2 + \tau_j T_3.$$

We compute the mean of the squared norm of $u_{j+1}^{h_{j+1}} - u^*$ as

$$\begin{aligned} (42) \quad \mathbb{E}[\|u_{j+1}^{h_{j+1}} - u^*\|^2] =& \mathbb{E}[\|u_j^{h_j} - u^*\|^2] + \tau_j^2 \mathbb{E}[\|T_1\|^2] + \tau_j^2 \mathbb{E}[\|T_2\|^2] + \tau_j^2 \mathbb{E}[\|T_3\|^2] \\ &- 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, T_1\rangle] + 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, T_2\rangle] + 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, T_3\rangle] \\ &- 2\tau_j^2 \mathbb{E}[\langle T_1, T_2\rangle] + 2\tau_j^2 \mathbb{E}[\langle T_2, T_3\rangle] - 2\tau_j^2 \mathbb{E}[\langle T_1, T_3\rangle]. \end{aligned}$$

Next, we will bound each of these ten terms to find a recursive formula on $\mathbb{E}[\|u_j^{h_j} - u^*\|^2]$. First, the term $\tau_j^2 \mathbb{E}[\|T_1\|^2]$ can be bounded as in the proof of Theorem 6 leading to:

$$\tau_j^2 \mathbb{E}[\|T_1\|^2] \leq \tau_j^2 L_{h_j}^2 \mathbb{E}[\|u_j^{h_j} - u^*\|^2],$$

with L_{h_j} being the Lipschitz constant for the function f^{h_j}, which is bounded by L (see Lemma 4). For the term $\tau_j^2 \mathbb{E}[\|T_3\|^2]$, we find,

$$\begin{aligned} \tau_j^2 \mathbb{E}[\|T_3\|^2] &= \tau_j^2 \|\mathbb{E}[\nabla f(u^*) - \nabla f^{h_j}(u^*)]\|^2 \\ &= \tau_j^2 \|\mathbb{E}[p(u^*) - p^{h_j}(u^*)]\|^2 \\ &\leq \tau_j^2 \mathbb{E}[\|p(u^*) - p^{h_j}(u^*)\|^2] \\ &\leq 2\tau_j^2 \mathbb{E}[\|p(u^*) - \widetilde{p}^{h_j}(u^*)\|^2] + 2\tau_j^2 \mathbb{E}[\|\widetilde{p}^{h_j}(u^*) - p^{h_j}(u^*)\|^2] \\ &\leq 2C\tau_j^2 \mathbb{E}[|p(u^*)|_{H^{r+1}}^2] h^{2r+2} + 2C\tau_j^2 \mathbb{E}[|y(u^*)|_{H^{r+1}}^2] h^{2r+2} \quad \text{[using Céa's Lemma]} \\ &\leq 2\tau_j^2 C(y(u^*), p(u^*)) h^{2r+2}. \end{aligned}$$

Next, for $\tau_j^2 \mathbb{E}[\|T_2\|^2]$ we use the same steps as in Theorem 6 to find

$$\tau_j^2 \mathbb{E}[\|T_2\|^2] \leq \frac{2\tau_j^2 L_{h_j}^2}{N_j} \mathbb{E}\Big[\|u_j^{h_j} - u^*\|^2\Big] + \frac{2\tau_j^2}{N_j} \mathbb{E}\Big[\|\nabla f^{h_j}(u^*, \omega)\|^2\Big].$$

Then we bound the second term of the right hand side uniformly w.r.t. h_j by

$$\begin{aligned} \|\nabla f^{h_j}(u^*, \omega)\|^2 &\leq 2\|\nabla f^{h_j}(u^*, \omega) - \nabla f(u^*, \omega)\|^2 + 2\|\nabla f(u^*, \omega)\|^2 \\ &\leq 4C(y(u^*), p(u^*)) h^{2r+2} + 2\|\nabla f(u^*, \omega)\|^2, \end{aligned}$$

where we have used the same steps as for T_3 to bound $\|\nabla f_{h_j}(u^*,\omega) - \nabla f(u^*,\omega)\|$. Finally, for the cross terms we have

$$\begin{aligned}
2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, T_1\rangle] &= 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, G[\nabla f^{h_j}(u_j^{h_j}) - \nabla f^{h_j}(u^*)]\rangle] \\
&= 2\tau_j \mathbb{E}[G[\langle u_j^{h_j} - u^*, \nabla f^{h_j}(u_j^{h_j}) - \nabla f^{h_j}(u^*)\rangle]] \quad \text{[using Strong convexity]} \\
&\geq \tau_j l \mathbb{E}[\|u_j^{h_j} - u^*\|^2],
\end{aligned}$$

and as in Theorem 9,

$$2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, T_2\rangle] = 2\tau_j^2 \mathbb{E}[\langle T_1, T_2\rangle] = 2\tau_j^2 \mathbb{E}[\langle T_2, T_3\rangle] = 0.$$

Moreover

$$\begin{aligned}
2\tau_j \mathbb{E}[\langle u_j^{h_j} - u^*, T_3\rangle] &\leq 2\tau_j \frac{l}{4} \mathbb{E}[\|u_j^{h_j} - u^*\|^2] + \frac{2\tau_j}{l} \mathbb{E}[\|T_3\|^2] \\
&\leq 2\tau_j \frac{l}{4} \mathbb{E}[\|u_j^{h_j} - u^*\|^2] + \frac{4\tau_j}{l} C(y(u^*), p(u^*)) h^{2r+2},
\end{aligned}$$

and finally

$$\begin{aligned}
2\tau_j^2 \mathbb{E}[\langle T_1, T_3\rangle] &\leq \tau_j^2 \mathbb{E}[\|T_1\|^2] + \tau_j^2 \mathbb{E}[\|T_3\|^2] \\
&\leq \tau_j^2 L_{h_j}^2 \mathbb{E}[\|u_j^{h_j} - u^*\|^2] + 2\tau_j^2 C(y(u^*), p(u^*)) h^{2r+2}.
\end{aligned}$$

Putting everything together, we finally obtain (41), as claimed. □

A natural choice to tune the parameters τ_j, N_j and h_j would be to set, guided by the usual Robbins-Monro theory, $\tau_j = \tau_0/j$, $N_j = \overline{N}$ and balancing all terms on right hand side of (41).

Theorem 9. *Suppose that the assumptions of Corollary 1 hold and let $u_j^{h_j}$ denote the j-th iterate of* (40). *For the particular choice $(\tau_j, N_j, h_j) = (\tau_0/j, \overline{N}, h_0 2^{-\ell(j)})$, with $\ell(j) = \lceil \frac{\ln_2(j) - \ln_2(\tau_0 l)}{2r+2} \rceil$, and assuming $\tau_0 > 1/l$, we have:*

$$\mathbb{E}[\|u_j^{h_j} - u^*\|^2] \leq F_1 j^{-1} \tag{43}$$

for a suitable constant F_1 independent of j.

Proof. With the choice of τ_j, N_j and $\ell(j)$ in the statement of the theorem, the two last terms $\frac{4\tau_j^2}{N_j}\mathbb{E}[\|\nabla f_{h_j}(u^*,\omega)\|^2]$ and $4\tau_j\big(\tau_j(1 + \frac{2}{N_j}) + \frac{1}{l}\big)Ch_j^{2r+2}$ in the inequality (41) have the same order $O(j^{-2})$. Then, we apply the same reasoning as in Theorem 7 to conclude the proof. □

Now we present the idea of the SG algorithm 3 with variable mesh size.

```
Algorithm 3: Stochastic Gradient with variable mesh size algorithm
Data:
Given a desired tolerance $tol$, choose $\frac{1}{l} < \tau_0$, $h_0$ and $j_{max} \simeq tol^{-2}$
initialization:
$u = 0$
for $j = 1, \dots, j_{max}$ do
    update mesh size to $h = h_0 2^{-\lceil \frac{\ln_2 j - \ln_2 \tau_0 l}{2r+2} \rceil}$
    sample one realization $a_j = a(\cdot, \omega_j)$ or the random field
    solve primal problem $\to y(a_j, u)$ on mesh $h$
    solve dual problem $\to p(a_j, u)$ on mesh $h$
    $\widehat{\nabla J} = \alpha u + p(a_j, u)$
    $u = u - \tau_j \widehat{\nabla J}$
end
```

Concerning the complexity of Algorithm 3, one can derive the following complexity result.

Corollary 4. *In order to achieve a given tolerance $O(tol)$, i.e. to guarantee that $\mathbb{E}[\|u_j^{h_j} - u^*\|^2] \lesssim tol^2$, the total required computational work W is bounded by:*

$$W \lesssim tol^{-2-\frac{n\gamma}{r+1}}$$

Proof. To achieve $tol^2 \lesssim j_{max}^{-1}$ requires $j_{max} \simeq tol^{-2}$. Then the total work required is bounded by

$$W = \sum_{p=1}^{j_{max}} 2\overline{N} h_p^{-n\gamma} = 2\overline{N} \sum_{p=1}^{j_{max}} 2^{n\gamma \lceil \frac{\ln_2 p - \ln_2 \tau_0 l}{2r+2} \rceil}$$

But as $\lceil \frac{\ln_2 p - \ln_2 \tau_0 l}{2r+2} \rceil \leq \frac{\ln_2 p - \ln_2 \tau_0 l}{2r+2} + 1$, one can bound:

$$\begin{aligned} W &\leq 2\overline{N} \sum_{p=1}^{j_{max}} 2^{n\gamma \left(\frac{\ln_2 p - \ln_2 \tau_0 l}{2r+2} + 1 \right)} \leq 2^{n\gamma+1} \overline{N} \{\tau_0 l\}^{\frac{-n\gamma}{2r+2}} \sum_{p=1}^{j_{max}} p^{\frac{n\gamma}{2r+2}} \\ &\leq 2^{n\gamma+1} \overline{N} \{\tau_0 l\}^{\frac{-n\gamma}{2r+2}} \frac{2r+2}{2r+2+n\gamma} (j_{max}+1)^{\frac{n\gamma}{2r+2}+1} \end{aligned}$$

But as $j_{max} \simeq tol^{-2}$, we finally bound the computational work by

$$W \lesssim tol^{-2-\frac{n\gamma}{r+1}}.$$

□

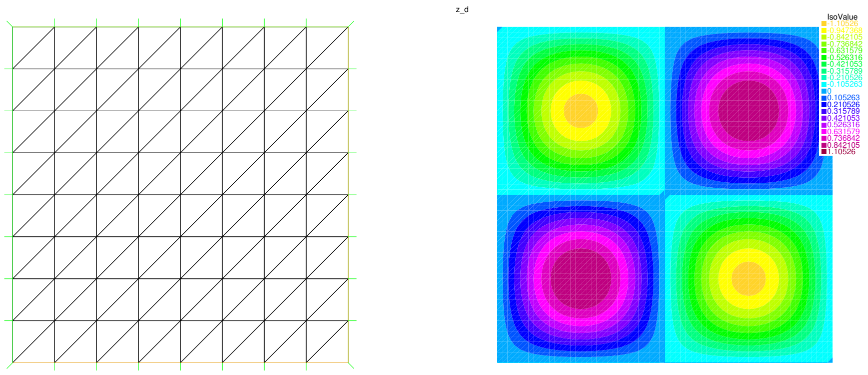
We notice that the asymptotic complexity remains the same as in the Stochastic Gradient algorithm with fixed mesh size. However, as we only use computations on coarse meshes for the first iterations, we thus expect an improvement due to reducing the constant. We will compute this constant, based on numerical examples, in the Section 8.

8. Numerical results

In this section we verify the assertions of Theorems 5, 8, and 9, as well as the computational complexity derived in the corresponding Corollaries. Specifically, we illustrate the order of convergence for the three versions of the steepest descent algorithm presented in Sections 5, 6, and 7 respectively. For this purpose, we consider the optimal control problem (19) with a MC approximation of the expectation. We consider problem (2) in the domain $D = (0,1)^2$ with $g = 1$ and the random diffusion coefficient

$$a(x_1, x_2, \boldsymbol{\xi}) = 1 + 0.1 \Big(\xi_1 \cos(\pi x_2) + \xi_2 \cos(\pi x_1) + \xi_3 \sin(2\pi x_2) + \xi_4 \sin(2\pi x_1) \Big), \tag{44}$$

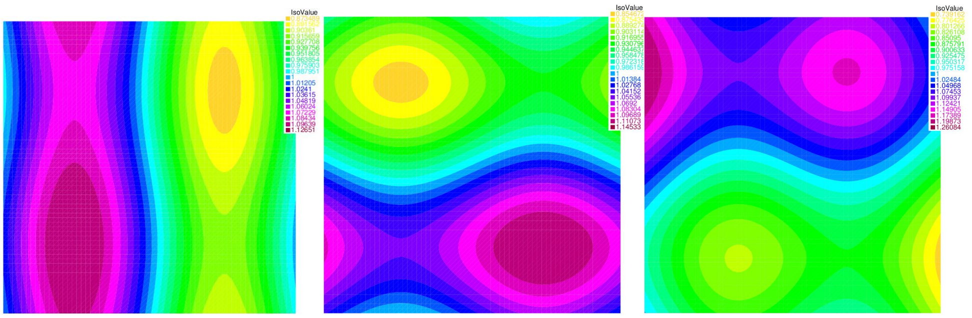
with $(x_1, x_2) \in D$ and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_4)$ with $\xi_i \overset{iid}{\sim} \mathcal{U}([-1,1])$. Figure 2 shows three typical realizations of the random field. The target function z_d has been chosen as $z_d(x,y) = \sin(2\pi x)\sin(2\pi y)$ (see Fig. 1 b) and we have taken $\alpha = 0.1$ in the objective function $J(u)$ in (3). For the FE approximation, we have considered a structured triangular grid of size h (see Fig. 1 a) where each side of the domain D is divided into $1/h$ sub-intervals and used piece-wise linear FE (i.e. $r = 1$). All calculations have been performed using the FE library Freefem++[20].



(A) Structured mesh triangulation with $h = 2^{-3}$

(B) Target function z_d for the optimal control problem

FIGURE 1. Mesh and target function z_d.



(A) $Y_1 = 0.0327973$
$Y_2 = 0.10508$
$Y_3 = 0.141335$
$Y_4 = 0.905369$

(B) $Y_1 = 0.370554$
$Y_2 = 0.0682218$
$Y_3 = 0.667794$
$Y_4 = -0.421315$

(C) $Y_1 = -0.943052$
$Y_2 = 0.968895$
$Y_3 = -0.656957$
$Y_4 = -0.997339$

FIGURE 2. Three realizations of the diffusion random field (44).

Reference solution. To compute a reference solution of problem (2), we use a full tensorized Gaussian Legendre (GL) quadrature grid with 5 points in each direction and a fine triangulation with $h = 2^{-8}$ (see, e.g., references [8, 37] and Appendix A.2 for a formal error estimate). As this approximated problem is now deterministic with fixed Gaussian nodes, we used a stopping condition on the gradient. In Figure 3 we show the optimal control obtained after $j = 6$ iterations when the stopping criterion $\|E^{GL}_{(5,5,5,5)}[\nabla J(u^h_j)]\| \leq 10^{-8}$ was met, where u^h_j is the j-th iterate of (19) and $\widehat{E}$ in (19) is a full tensorized Gaussian Legendre (GL) quadrature approximation of the expectation. The steepest descent step size was chosen as $\tau_0 = 10$. The L^2-norm of the final control using this Gaussian quadrature is $\|\widehat{u}^{h=2^{-8}}_{j=6}\| = 0.0663345$.

8.1. **Steepest descent algorithm with fixed discretization.** We investigate here the convergence of the method defined in (24), for which we recall the error bound (30) in the case of piece-wise linear FE (i.e. $r = 1$):

$$\mathbb{E}[\|\widehat{u}^h_j - u\|^2] \leq C_1 e^{-\rho j} + \frac{C_2}{N} + C_3 h^4 \ .$$

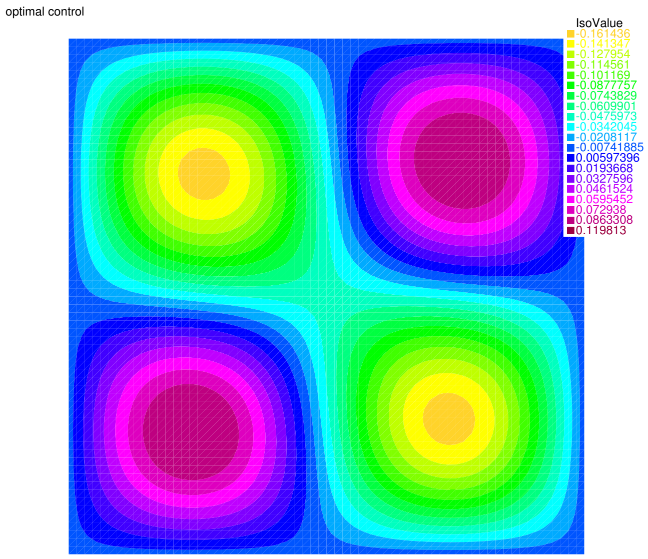


FIGURE 3. Optimal control reference solution computed with $h = 2^{-7}$ on tensorized Gauss-Legendre quadrature formula with $N = 9^4$ nodes.

For each tolerance tol, using formula (30), we compute the optimal mesh size $h = f_1(tol)$, the optimal sample size in the MC approximation $N = f_2(tol)$, and, finally, the minimum number of iterations we need in the steepest descend method, $j_{max} = f_3(tol)$. Here, the three functions f_i are introduced to emphasize that these parameters are completely determined by the prescribed tolerance goal tol. In what follows, we compare the actual error on the optimal control obtained from the algorithm (measured w.r.t. the reference solution) with the prescribed tolerance.

Estimation of the constants C_1, C_2, C_3. To have a precise idea of the functions f_i, we have estimated the constants in (30) numerically.

- In order to estimate C_1 we used the same finest mesh as the one used to compute our reference solution with $h = 2^{-8}$, and we used also the same Gaussian 5 points for the quadrature. We computed numerically the squared error between the optimal control after i iterations and the reference solution computed above. We then only see the first term in (30), and running the algorithm for the first 10 iterations of the steepest descent method, we estimated a constant $C_1 \approx 10^{-3}$ and $\rho \approx 3.2$.
- To estimate the second constant C_2, we used again the same finest mesh as the one used to compute our reference solution with $h = 2^{-8}$. We ran the steepest descent method up to 10 iterations, using a MC estimator for the mean of the gradient with a sample size N_{MC} of $N_{MC} = 2^0, 2^1, \cdots, 2^5$. Finally, for every sample size N_{MC} of the MC estimator, we averaged the final error squared on the control over 10 independent realizations. As we go up to 10 iterations, the error term is of order $C_1 e^{-3.2\times 10} = 1.27 \times 10^{-17}$. That is, as long as the term C_2/N stays bigger than 10^{-15}, i.e. $C_2 > 10^{-14}$, we effectively only see the C_2/N term. We numerically found $C_2 \approx 3.16 \times 10^{-5}$, what is coherent with the last condition.
- Finally, to compute the third term, we used different mesh sizes $h = 2^{-1}, \cdots, 2^{-5}$, and we used a steepest descent algorithm with sufficiently

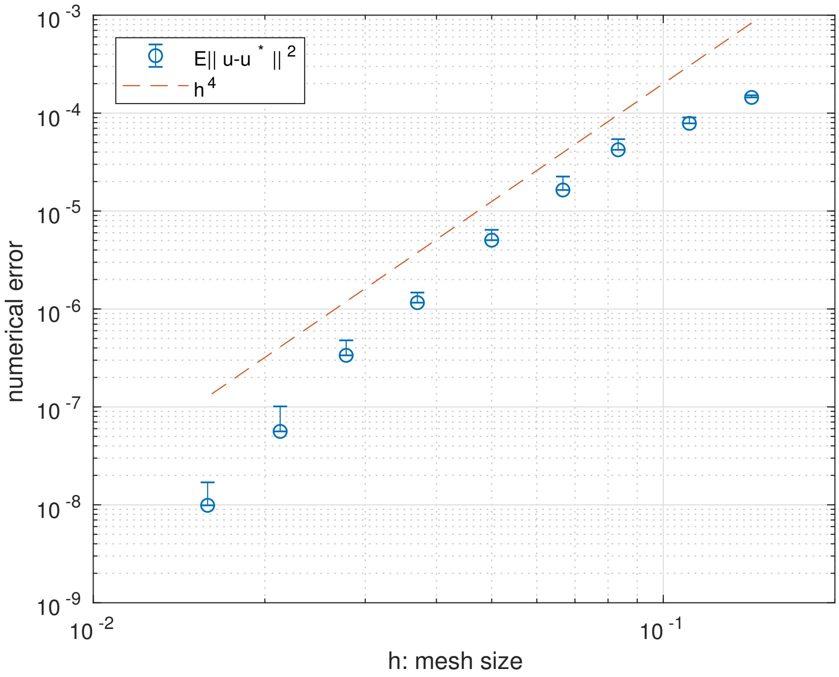


Figure 4. Steepest descent Algorithm 1 with fixed discretization over iterations. Error $\mathbb{E}[\|u-u^*\|^2]$ as a function of the mesh size h. Blue circles: estimated mean over 20 repetitions. Error bars: one estimated standard deviation.

many iterations with a Gaussian quadrature with 5 points in each direction. We found $C_3 \approx 5.01 \times 10^{-1}$.

Figure 4 shows the convergence of the error on the control (in the L^2-norm), versus the discretization parameter h (that is directly linked to N and j_{max} using the functions f_i, $i = 1, 2, 3$). The bars denote plus one standard deviation, estimated by repeating the simulation 20 times. We observe a convergence rate of h^{-4} on the squared error, which is consistent with the theoretical result (30). Figure 5 shows the corresponding computational complexity. Here we have used the theoretical computational cost $W = 2N j_{max} h^{-2}$ (which assumes an optimal linear algebra solver with $\gamma = 1$). The observed slope is consistent with our theoretical result $W \sim tol^{-3}$ up to logarithmic terms.

8.2. **Stochastic Gradient with fixed mesh size h.** We implemented here the Stochastic Gradient method described in Section 6 using $\overline{N} = 1$ sample at each iteration (recall that the complexity does not depend on $\overline{N}$). As the error result (35) is in the mean squared sense, we ran the simulation 10 times and averaged the obtained errors, in order to estimate this mean.

Estimation of the constants D_1, D_2. Also for the SG method with a fixed mesh size we have estimated the constants in (35).

- To numerically estimate the constant D_1, we simply used the finest mesh of size $h = 2^{-8}$ and plotted the squared error on the control versus the i-th iteration using a Stochastic Gradient technique. We repeated the procedure

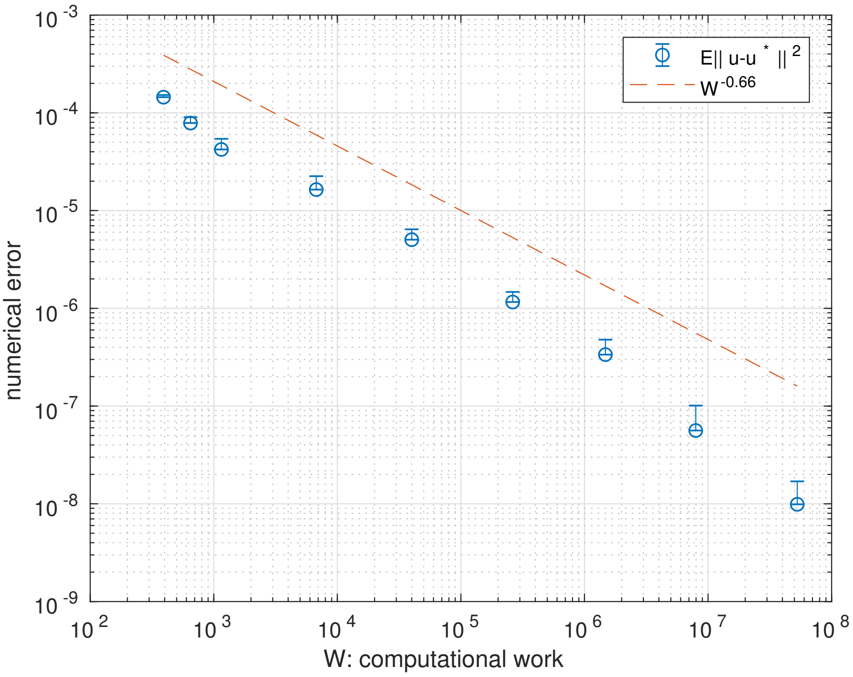


FIGURE 5. Steepest descent Algorithm 1 with fixed discretization over iterations. Error $\mathbb{E}[\|u-u^*\|^2]$ as a function of the theoretical computational work W. Blue circles: estimated mean over 20 repetitions (only 2 repetition in the last two points). Error bars: one estimated standard deviation.

10 times to compute a MC estimator of the expectation of this squared error. We found effectively a slope of -1 and the constant $D_1 \approx 2.51 \times 10^{-6}$.

- Again as for the fixed MC procedure, to estimate the second term constant D_2, we used different mesh sizes $h = 2^{-1}, \cdots, 2^{-5}$, and a Stochastic Gradient algorithm with sufficiently many iterations. We found $D_2 \approx 6.31 \times 10^{-1}$, which is very close to the C_3 constant, estimated earlier.

Figures 6 presents the squared error on the control for different desired tolerances tol, i.e. different mesh sizes, using the SG steepest descent method with resampling. The theoretical rate is thus verified for $r = 1$ and $d = 2$. Figure 7 and 8 show the estimated mean squared error, using Algorithm 2, as a function of the theoretical cost $W = 2j_{max}h^{-2}$. The slope is the one expected, namely $W \lesssim tol^{-3}$.

8.3. **Stochastic Gradient with decreasing mesh size h_j.** We illustrate here the Stochastic Gradient method described in Section 7. As the error result (43) is in mean-squared sense, we ran the simulation 20 times up to iteration $j_{max} = 4000$. We then average every error at every iteration over these 20 simulation. In Figure 9 we plot the averaged errors obtained versus the iteration of the SG recursion. In fact, the plot shows the mean squared errors and the mean squared errors plus one standard deviation, both obtained using once more all the 20 simulations. As we refine using only embedded mesh, we do see a refinement drop at iterations $j = 16, 256, 4096$. Notice that the next refinement would be at iteration $j = 65536$, which however is computationally prohibitive.

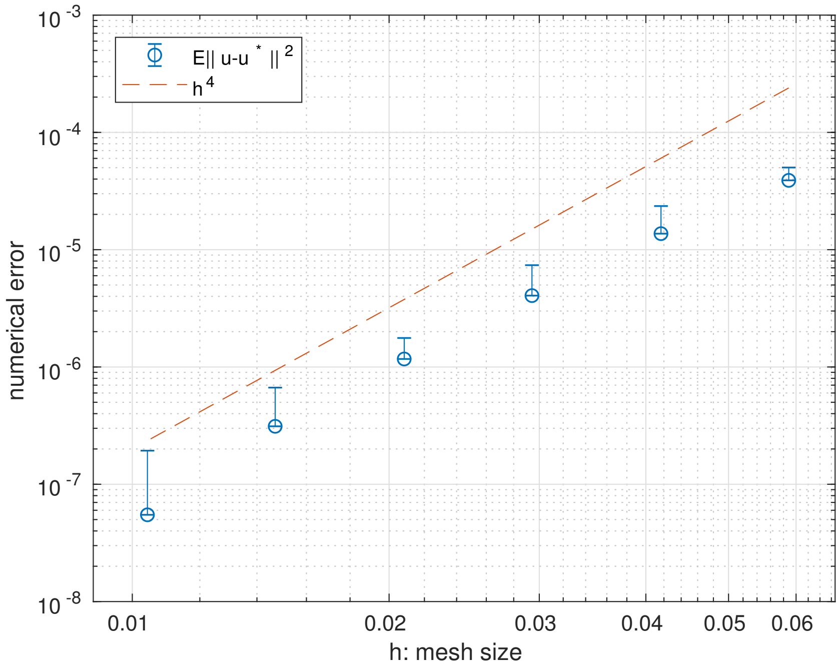


FIGURE 6. SG Algorithm 2 with fixed space discretization over iterations. Error $\mathbb{E}[\|u-u^*\|^2]$ as a function of the mesh size h. Blue circles: estimated mean over 10 repetitions (only 2 repetitions in the last two points). Error bars: one estimated standard deviation.

Estimation of the initial mesh size h_0. In practice, in order to estimate the parameter h_0, we set a desired final tolerance tol, which is directly related to h_{final} through the constants D_1 and D_2 estimated previously. Based on j_{max} linked to the tolerance tol and expression (43), we can thus determine the initial mesh size h_0. That is, with the initial mesh size h_0 fixed, we then run the algorithm with this h_0, ensuring that the algorithm will terminate at iteration j_{max} with final mesh size $h_{j_{max}}$.

In Figure 10 we plot the averaged numerical error versus the computational cost W for the three algorithm studied in the previous Sections: the fixed MC gradient, the SG with fixed mesh, and the SG with variable mesh size. For the fixed MC gradient and the SG with fixed mesh, we ran 20 iid simulations for every tolerance (i.e. every point and every square in the Figure) and then averaged them to estimate the mean. For the SG with variable mesh size we show 3 different realization of error versus computational work (with the same initial mesh size h_0). As discussed before, the SG is more efficient than the fixed MC gradient, but only by a logarithmic factor (which is difficult to observe in Figure 10). All three algorithms follow a slope of $tol^2 \sim W^{-2/3}$, as predicted by our theoretical complexity analysis. The proportionality constant is smaller for the SG compared to the fixed MC gradient, and seems to further reduce for the variable mesh size SG version at least in the range of computational works considered. This is consistent with our intuition that computational work is saved in the earlier iterations in this version of the SG method.

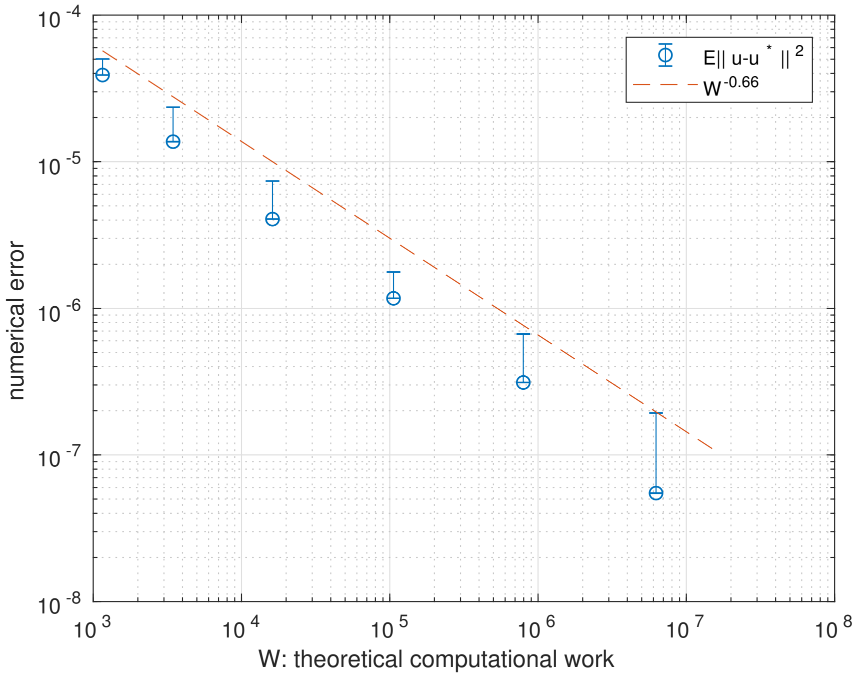


FIGURE 7. SG Algorithm 2 with fixed space discretization over iterations. The error $\mathbb{E}[\|u - u^*\|^2]$ as a function of the theoretical computational work W is plotted. Blue circles: estimated mean over 10 repetitions (only 2 repetitions in the last two points). Error bars: one estimated standard deviation.

9. Conclusions

In this work, we have analyzed and compared the complexity of three versions of the gradient method for the numerical solution of a mean-based risk-averse optimal control problem for an elliptic PDE with random coefficients, where a Finite Element discretization is used to approximate the underlying PDEs and a Monte Carlo sampling is used to approximate the expectation in the risk measure. In the first version the FE mesh and Monte Carlo sample are chosen initially and kept fixed over the iterations. In the second version, a Stochastic Gradient method, the finite element discretization is still kept fixed over the iterations, however the expectation in the objective function is re-sampled independently at each iteration, with a small (fixed) sample size. Finally, the third version is again a stochastic gradient method, but now with successively refined FE meshes over the iterations. We have shown in particular, that the stochastic versions of the gradient method improve the computational complexity by log factors. Our complexity analysis is based on a priori error estimates and a priori choices of the FE mesh size, the Monte Carlo sample size, and the gradient iterations to obtain a prescribed tolerance.

Beside the improved complexity, another interest in looking at stochastic versions of the gradient method is that they are more amenable to adaptive versions, in which the mesh size and possibly the Monte Carlo sample size are refined over the iterations based on suitable a posteriori error indicators. The study of such adaptive versions is postponed to future work.

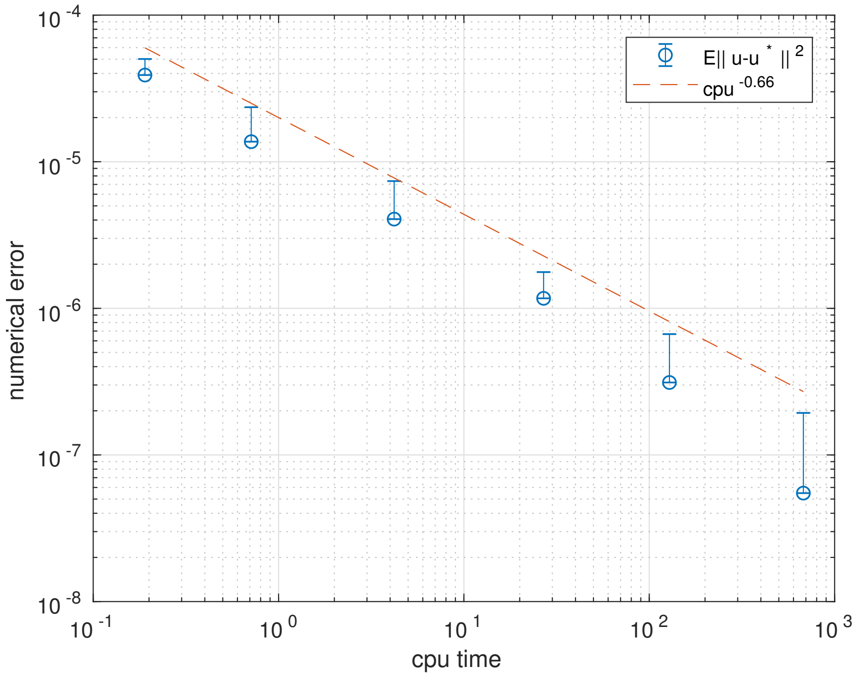


FIGURE 8. SG Algorithm 2 with fixed space discretization over iterations. The error $\mathbb{E}[\|u-u^{*}\|^{2}]$ as a function of the average CPU time is plotted. Blue circles: estimated mean over 10 repetitions (only 2 repetition in the last two points). Error bars: one estimated standard deviation.

Another interesting direction is the extension of stochastic gradient methods to more general risk measures. We mention that Stochastic Gradient methods have been already used in combination with the CVaR risk measure [5], although not in the context of PDE-constrained optimal control problems.

Acknowledgments

F. Nobile, S. Krumscheid, and M. Martin have received support from the Center for ADvanced MOdeling Science (CADMOS).

Appendix A. Reference solution by Stochastic Collocation

A.1. **Optimal Control Problem with quadrature.** In this appendix, we describe the computation of the reference solution used in the numerical result of Section 8, by the Stochastic Collocation method on a tensor grid of Gauss Legendre points and provide an error estimate for such reference solution. In the setting of Section 8, with only 4 random variables, we show here that the Stochastic Collocation approximation is exponentially convergent and a very accurate solution can be obtained with a moderate number of collocation points (5^{4} were used in the numerical results). We suppose here that our expectation estimator is not random, but uses deterministic points ξ_i, for $i = 1, \ldots, N$. The estimated optimal control $\widehat{u}$ is then deterministic as well. The following theorem derives an error bound when we estimate the exact expectation $\mathbb{E}$ in (3) by a deterministic quadrature formula $\widehat{E}$.

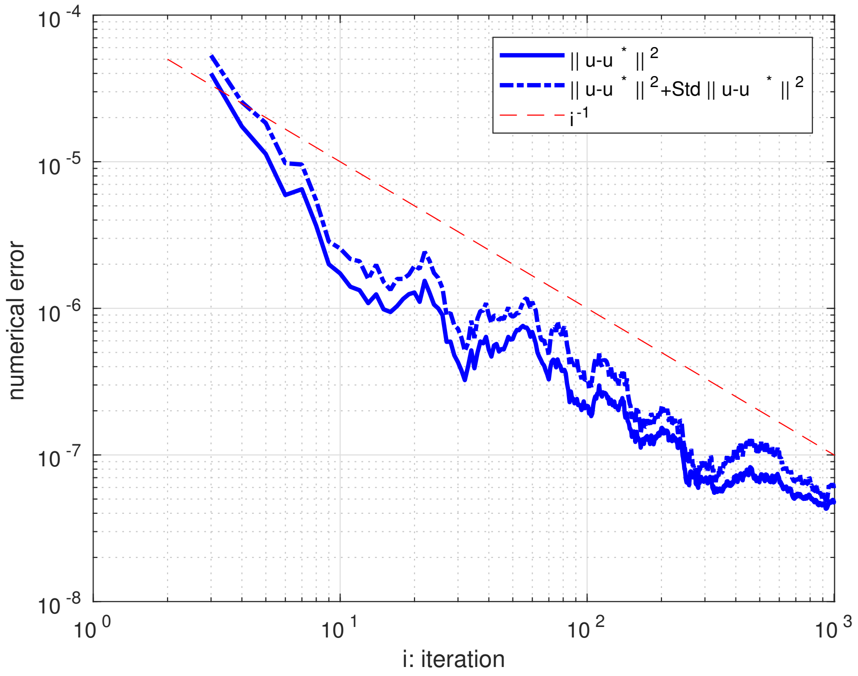


FIGURE 9. SG Algorithm 3 with variable mesh size over iterations. The error $\mathbb{E}[\|u-u^*\|^2]$ as a function of iteration index is plotted. Solid (blue) line: estimated mean over 20 repetitions. Dashed (blue) line: estimated mean plus one standard deviation.

Theorem 10. *Denoting by u^* the optimal control solution of the exact problem* (10) *and by $\widehat{u}$ the solution of the semi-discrete collocation problem* (19)*, we have*

$$\frac{\alpha}{2}\|\widehat{u}-u^*\|^2+\mathbb{E}[\|y(u^*)-y(\widehat{u})\|^2] \leq \frac{1}{2\alpha}\|\mathbb{E}[p(\widehat{u})]-\widehat{E}[p(\widehat{u})]\|^2. \tag{45}$$

Proof. The expressions of the gradient of J and $\widehat{J}$ are given by $\nabla J(u^*)=\alpha u^*+\mathbb{E}[p(u^*)]$, $\nabla\widehat{J}(\widehat{u})=\alpha\widehat{u}+\widehat{E}[p(\widehat{u})]$. From the optimality condition (5) for J, we derive the optimality condition for $\widehat{J}$ as:

$$\langle\nabla\widehat{J}(\widehat{u}), v'-\widehat{u}\rangle \geq 0 \quad \forall v' \in U. \tag{46}$$

Then choosing $v=\widehat{u}$ in (5) and $v'=u^*$ in (46) and combining both, we have

$$\langle\alpha(u^*-\widehat{u})+\mathbb{E}[p(u^*)]-\widehat{E}[p(\widehat{u})], \widehat{u}-u^*\rangle \geq 0,$$

that is,

$$\alpha\|u^*-\widehat{u}\|^2 \leq \langle\mathbb{E}[p(u^*)]-\mathbb{E}[p(\widehat{u})]+\mathbb{E}[p(\widehat{u})]-\widehat{E}[p(\widehat{u})], \widehat{u}-u^*\rangle. \tag{47}$$

In order to bound the first part of the error in (47), $\langle\mathbb{E}[p(u)]-\mathbb{E}[p(\widehat{u})], \widehat{u}-u\rangle$, we take one random realization ω and we use the primal-dual equations to obtain:

$$\begin{aligned}\langle\widehat{u}-u^*, p_\omega(u^*)-p_\omega(\widehat{u})\rangle &= b_\omega(y_\omega(\widehat{u})-y_\omega(u^*), p_\omega(u)-p_\omega(\widehat{u}))\\ &= \langle y_\omega(u^*)-y_\omega(\widehat{u}), y_\omega(\widehat{u})-y_\omega(u^*)\rangle\\ &= -\|y_\omega(u^*)-y_\omega(\widehat{u})\|^2.\end{aligned}$$

Then taking the (exact) expectation over all the realizations ω, we find:

$$\langle\mathbb{E}[p(u^*)]-\mathbb{E}[p(\widehat{u})], \widehat{u}-u^*\rangle = -\mathbb{E}[\|y(u^*)-y(\widehat{u})\|^2].$$

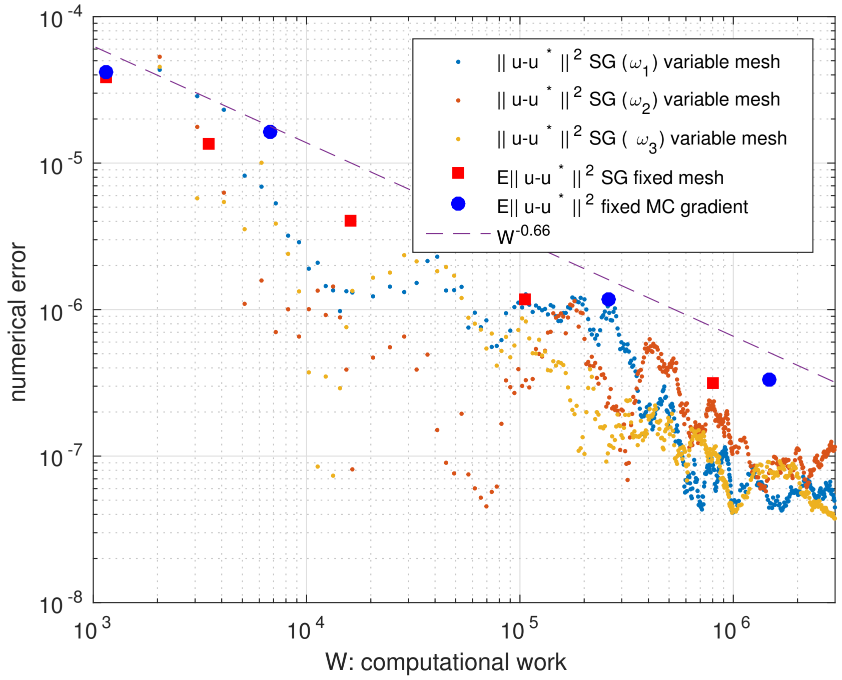


FIGURE 10. Comparison between Algorithm 1, 2, 3. The estimated mean squared error $\mathbb{E}[\|u-u^*\|^2]$ is plotted as a function of the theoretical computation work W for Fixed MC gradient and SG with fixed mesh, versus 3 different realization of the SG with variable mesh size algorithm.

For the second contribution, $\langle \mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})], \widehat{u} - u^* \rangle$, we simply use Young's inequality, yielding

$$\langle \mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})], \widehat{u} - u^* \rangle \leq \frac{1}{2\alpha} \|\mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})]\|^2 + \frac{\alpha}{2}\|\widehat{u} - u^*\|^2,$$

from which the claim eventually follows. □

A.2. **Collocation on tensor grid of Gauss points (Gaussian Legendre quadrature).** The quantification of the quadrature error $\mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})]$, i.e. the right hand side in (45), heavily depends on the smoothness of the dual function in the stochastic variables. The numerical example considered in Section 8 has a diffusion coefficient of the form

$$a(x,\xi) = a_0(x) + \sum_{i=1}^{M} \sqrt{\lambda_i}\xi_i b_i(x) \ ,$$

with $a_0 > 0$ a.e. in D, $\|b_i\|_{L^\infty(D)} = 1$, $\sum_{i=1}^{M} \sqrt{\lambda_i} < \operatorname{essinf}_{x \in D} a_0(x)$ and $\xi_i \sim \mathcal{U}([-1,1])$ iid uniform random variables. We denote by $\xi = (\xi_1, \cdots, \xi_M)$ the corresponding random vector. Hence, in this case the probability space $(\Gamma, \mathcal{F}, P)$ is $\Gamma = [-1,1]^M$, $\mathcal{F} = \mathcal{B}(\Gamma)$ the Borel σ-algebra on Γ, and $\mathbb{P}(d\xi) = \otimes_{i=1}^{M} \frac{d\xi_i}{2}$ the uniform product measure on Γ. In this case we chose as a quadrature formula the tensor Gaussian quadrature built on Gauss-Legendre quadrature points. In particular, we consider a tensor grid with q_i points in the i-th variable and denote the corresponding quadrature by $E_q^{GL}[\cdot]$, where $q = (q_1, \cdots, q_M) \in \mathbb{N}^M$ is a multi-index.

To any vector of indexes $[k_1, \dots, k_M] \in \prod_{i=1}^{M} \{1, \cdots, q_i\}$ we associate the global index

$$k = k_1 + q_1(k_2 - 1) + q_1 q_2(k_3 - 1) + \dots,$$

and we denote by y_k the point $y_k = [y_{1,k_1}, y_{2,k_2}, ..., y_{M,k_M}] \in \Gamma$. We also introduce, for each $n = 1, 2, \dots, N$, the Lagrange basis $\{l_{n,j}\}_{j=1}^{q_n}$ of the space $P_{q_n - 1}$,

$$l_{n,j} \in P_{q_n - 1}(\Gamma_n), \quad l_{n,j}(y_{n,k}) = \delta_{jk}, \quad j, k = 1, \dots, q_n,$$

where δ_{jk} is the Kronecker symbol, and $P_{q-1}(\Gamma) \subset L^2(\Gamma)$ is the span of tensor product polynomials with degree at most $q - 1 = (q_1 - 1, \dots, q_M - 1)$; i.e., $P_{q-1}(\Gamma) = \bigotimes_{i=1}^{M} P_{q_i - 1}(\Gamma_i)$. Hence the dimension of P_{q-1} is $N_q = \prod_{i=1}^{N}(q_i)$. Finally we set $l_k(y) = \prod_{n=1}^{N} l_{n,k_n}(y_n)$.

For any continuous function $g : \Gamma \to \mathbb{R}$ we introduce the Gauss Legendre quadrature formula $E_q^{GL}[g]$ approximating the integral $\int_\Gamma g(y)\, \mathrm{d}y$ as

$$E_q^{GL}[g] = \sum_{k=1}^{N_q} \omega_k g(y_k), \quad \omega_k = \prod_{n=1}^{M} \omega_{k_n}, \quad \omega_{k_n} = \int_{\Gamma_n} l_{k_n}^2(y)\, \mathrm{d}y \tag{48}$$

We now analyze the error introduced by the quadrature formula. The first step is to investigate the smoothness of the map $\xi \mapsto p(\widehat{u}, \xi)$. For this, it is convenient to extend the primal and dual problems to the complex domain. To do so, let us define

$$a(x, z) = a_0(x) + \sum_{i=1}^{M} \sqrt{\lambda_i} z_i b_i(x)$$

with $z = (z_1, \cdots, z_M) \in \mathbb{C}^M$ and let

$$\mathcal{U}_0 = \{z \in \mathbb{C}^M : \mathcal{R}e(a(x, z)) > 0 \quad a.e. \quad \text{in} \quad D\}.$$

We consider the primal and dual problems extended to the complex domain: $\forall z \in \mathcal{U}_0$ find $y(\cdot, z) \in H_0^1(D; \mathbb{C})$ s.t.

$$\int_D a(x, z) \nabla y(x, z) \nabla v(x) \mathrm{d}x = \int_D (\widehat{u}(x) + g(x)) v(x) \mathrm{d}x \quad \forall v \in H_0^1(D; \mathbb{C}) \ , \tag{49}$$

and find $p(\cdot, z) \in H_0^1(D; \mathbb{C})$ s.t.

$$\int_D a(x, z) \nabla p(x, z) \nabla v(x) \mathrm{d}x = \int_D (y(x, z) - z_d(x)) v(x) \mathrm{d}x \quad \forall v \in H_0^1(D; \mathbb{C}) \ . \tag{50}$$

It is well known that problem (49) and (50) are well posed in $\mathcal{U}_0$. Let now $\Sigma \subset \mathcal{U}_0$ be

$$\Sigma := \{z \in \mathbb{C}^N : \sum_{i=1}^{M} \sqrt{\lambda_i} |z_i| \leq \frac{a_{min}}{2}\}$$

with $a_{min} = ess \inf_{x \in D} a_0(x)$. The next Lemma states that both $z \mapsto y(\cdot, z)$ and $z \mapsto p(\cdot, z)$ are holomorphic functions in $\mathcal{U}_0$ with uniform bounds on Σ. The result for $z \mapsto y(\cdot, z)$ is well known and can be found in reference [11] fro example, so that we only give the proof for $z \mapsto p(\cdot, z)$.

Lemma 7. *Both functions $z \mapsto y(\cdot, z)$ and $z \mapsto p(\cdot, z)$ are holomorphic on $\mathcal{U}_0$, and both have a uniform bound on Σ, in the sense that*

$$\max_{z \in \Sigma} \|y(\cdot, \xi)\|_{H_0^1} \leq C_P \frac{\|g + \widehat{u}\|}{a_{min}} \tag{51}$$

and

$$\max_{z \in \Sigma} \|p(\cdot, z)\|_{H_0^1} \leq C_P \frac{\|z_d\|}{a_{min}} + C_P^3 \frac{\|g + \widehat{u}\|}{a_{min}^2} \ . \tag{52}$$

Proof. It is well known (see e.g. [11]) that the function $z \mapsto y(\cdot, z)$ is holomorphic on $\mathcal{U}_0$ with bound (51). This property translates to the dual function $z \mapsto p(\cdot, z) \in H_0^1(D; \mathbb{C})$ which is holomorphic in $\mathcal{U}_0$ as well with bound

$$\begin{aligned}\max_{z\in\Sigma} \|p(\cdot, z)\|_{H^1} &\leq C_P \max_{z\in\Sigma} \frac{\|y(\cdot, z) - z_d\|}{a_{min}} \\ &\leq C_P \frac{\|z_d\|}{a_{min}} + C_P \max_{z\in\Sigma} \frac{\|y(\cdot, z)\|}{a_{min}} \\ &\leq C_P \frac{\|z_d\|}{a_{min}} + C_P^3 \frac{\|g + \widehat{u}\|}{a_{min}^2} \,.\end{aligned}$$

□

Based on the last regularity result and following [3], we can state the following error estimate for the quadrature error.

Theorem 11. *Denoting by $\widehat{u}$ the solution of the semi-discrete (in probability) optimal control problem (19) with $\widehat{E} = E_q^{GL}[\cdot]$ and $p(\widehat{u})$ the corresponding adjoint function, there exists $C > 0$ and $\{r_1, \cdots, r_M\}$ independent of q s.t.*

$$\|\mathbb{E}[p(\widehat{u})] - E_q^{GL}[p(\widehat{u})]\|^2 \leq C \sum_{n=1}^{M} e^{-r_n q_n} \,,$$

with q_n the number of points used in the quadrature in direction n.

Appendix B. Proof for increasing Monte Carlo sampling in SG

Here we detail the proof of the bound (39) in remark 5. The factor c_j in (33) becomes

$$c_j := 1 - \tau_j l + L^2 \Big(1 + \frac{2}{N_j}\Big)\tau_j^2 = 1 - \frac{\tau_0 l}{j} + L^2 \Big(1 + 2j^{1-\tau_0 l}\Big)\frac{\tau_0^2}{j^2} \,,$$

for $\tau_j = \tau_0/j$ and $N_j \sim i^{\tau_0 l - 1}$ with $\tau_0 l - 1 > 0$. We use the recursive formula (33) and set, as before, u^{h*} to be the exact optimal control for the FE problem defined in (11). We emphasize that (11) has no approximation in probability space. Setting $a_j = \mathbb{E}[\|u_j^h - u^{h*}\|^2]$ and $\beta_j = \frac{2\tau_j^2}{N_j}\mathbb{E}[\|\nabla f(u^{h*}, \omega)\|^2]$, we have from (33), applied to the sequence of FE solutions $\{u_j^h\}_{j>0}$,

$$\begin{aligned}a_{j+1} \leq& c_j a_j + \beta_j \\ \leq& c_j c_{j-1} a_{j-1} + c_j \beta_{j-1} + \beta_j \\ \leq& \cdots \\ \leq& \underbrace{\Big(\prod_{i=1}^{j} c_i\Big)}_{=\kappa_j} a_1 + \underbrace{\sum_{i=1}^{j} \beta_i \prod_{l=i+1}^{j} c_l}_{=\mathcal{B}_j} \,.\end{aligned} \tag{53}$$

For the first term κ_j, computing its logarithm, we have

$$\log(\kappa_j) \leq \sum_{i=1}^{j} \log(1 - \frac{\tau_0 l}{i} + \frac{M'}{i^2}) \leq \sum_{i=1}^{j} \frac{-\tau_0 l}{i} + \sum_{i=1}^{j} \frac{M'}{i^2},$$

where we have set $M' = 3\tau_0^2 L^2$ as we have $1 - \tau_0 l < 0$ and thus $j^{1-\tau_0 l} \leq 1$ for every $j \geq 1$. Therefore

$$\log(\kappa_j) \leq -\tau_0 l \log j + M'', \text{ with } M'' = \sum_{i=1}^{\infty} \frac{M'}{i^2}$$

and $\kappa_j \lesssim j^{-\tau_0 l}$. For the second term $\mathcal{B}_j$ in (53) we have

$$\mathcal{B}_j = \sum_{i=1}^{j} \beta_i \prod_{k=i+1}^{j} c_k \le \sum_{i=1}^{j} S' i^{-\tau_0 l - 1} \underbrace{\prod_{k=i+1}^{j} \Big(1 - \frac{\tau_0 l}{k} + \frac{3\tau_0^2 L^2}{k^2}\Big)}_{=K_{ij}}, \quad \text{with } S' = 2\tau_0^2 \mathbb{E}[\|\nabla f(u^{h*}, \omega)\|^2] \,.$$

For the term K_{ij} we find that

$$\begin{aligned}\log(K_{ij}) &= \sum_{k=i+1}^{j} \log\left(1 - \frac{\tau_0 l}{k} + \frac{M'}{k^2}\right) \\ &\le \sum_{k=i+1}^{j} \Big(-\frac{\tau_0 l}{k} + \frac{M'}{k^2}\Big) \\ &\le -\tau_0 l(\log(j+1) - \log(i+1)) + M'\left(\frac{1}{i} - \frac{1}{j}\right),\end{aligned}$$

which shows that

$$K_{ij} \le (j+1)^{-\tau_0 l}(i+1)^{\tau_0 l} \exp\left(M'\left(\frac{1}{i} - \frac{1}{j}\right)\right) \,.$$

It follows that

$$\begin{aligned}\mathcal{B}_j &\le (j+1)^{-\tau_0 l} \underbrace{\exp\left(-\frac{M'}{j}\right)}_{\le 1} \sum_{i=1}^{j} S' i^{-\tau_0 l - 1}(i+1)^{\tau_0 l} \underbrace{\exp\left(\frac{M'}{i}\right)}_{\le \exp(M')} \\ &\le S' \exp(M')(j+1)^{-\tau_0 l} \sum_{i=1}^{j} (i+1)^{-1} \lesssim j^{-\tau_0 l} \log(j),\end{aligned}$$

for $\tau_0 > 1/l$. Eventually, we obtained the following upper bound for two constants $D_3 > 0$ and $D_4 > 0$:

$$a_{j+1} \le D_3 j^{-\tau_0 l} a_1 + D_4 j^{-\tau_0 l} \log(j). \tag{54}$$

We conclude that

$$a_{j+1} \le D_4 j^{-\tau_0 l} \log(j), \tag{55}$$

with D_4 possibly depending on $\|u_0^h - u^{h*}\|$. Finally, splitting the error as

$$\mathbb{E}[\|u_j^h - u^*\|^2] \le 2\mathbb{E}[\|u_j^h - u^{h*}\|^2] + 2\mathbb{E}[\|u^{h*} - u^*\|^2]$$

and using (18) to bound the second term, the claim follows.

References

[1] A. Ahmad Ali, E. Ullmann, and M. Hinze, *Multilevel Monte Carlo analysis for optimal control of elliptic PDEs with random coefficients*, SIAM/ASA J. Uncertain. Quantif. **5** (2017), no. 1, 466–492. MR 3640630

[2] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, *Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations*, SIAM/ASA J. Uncertain. Quantif. **5** (2017), no. 1, 1166–1192. MR 3725286

[3] I. Babuška, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM review **52** (2010), no. 2, 317–355.

[4] I. Babuska, R. Tempone, and G.E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM Journal on Numerical Analysis **42** (2004), no. 2, 800–825.

[5] O. Bardou, N. Frikha, and G. Pagès, *Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling*, Monte Carlo Methods Appl. **15** (2009), no. 3, 173–210. MR 2573212

[6] P. Benner, A. Onwunta, and M. Stoll, *Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs*, SIAM J. Matrix Anal. Appl. **37** (2016), no. 2, 491–518. MR 3483160

[7] A. Borzì and V. Schulz, *Computational optimization of systems governed by partial differential equations*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2012.

[8] A. Borzì, V. Schulz, C. Schillings, and G. von Winckel, *On the treatment of distributed uncertainties in PDE-constrained optimization*, GAMM-Mitteilungen **33** (2010), no. 2, 230–246.

[9] A. Borzì and G. von Winckel, *A POD framework to determine robust controls in PDE optimization*, Comput. Vis. Sci. **14** (2011), no. 3, 91–103. MR 2863632

[10] P. Chen and A. Quarteroni, *Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint*, SIAM/ASA J. Uncertain. Quantif. **2** (2014), no. 1, 364–396. MR 3283913

[11] A. Cohen and R. DeVore, *Approximation of high dimensional parametric PDEs*, Acta Numerica **24** (2015), 1–159.

[12] J. C. De los Reyes, *Numerical PDE-constrained optimization*, Springer, Cham, 2015.

[13] A. Défossez and F. Bach, *Averaged least-mean-squares: Bias-variance tradeoffs and optimal sampling distributions*, Artificial Intelligence and Statistics, 2015, pp. 205–213.

[14] A. Dieuleveut and F. Bach, *Nonparametric stochastic approximation with large step-sizes*, The Annals of Statistics **44** (2016), no. 4, 1363–1399.

[15] L.C. Evans, *Partial differential equations*, Graduate studies in mathematics, American Mathematical Society, 1998.

[16] N. Flammarion and F. Bach, *From averaging to acceleration, there is only a step-size*, Conference on Learning Theory, 2015, pp. 658–695.

[17] M. D. Gunzburger, H.-C. Lee, and J. Lee, *Error estimates of stochastic optimal Neumann boundary control problems*, SIAM J. Numer. Anal. **49** (2011), no. 4, 1532–1552. MR 2831060

[18] M.D. Gunzburger, C.G. Webster, and G. Zhang, *Stochastic finite element methods for partial differential equations with random input data*, Acta Numerica **23** (2014), 521–650.

[19] S. B. Hazra, *Large-scale PDE-constrained optimization in applications*, Springer-Verlag, Berlin, 2010.

[20] F. Hecht, *New development in freefem++*, J. Numer. Math. **20** (2012), no. 3-4, 251–265. MR 3043640

[21] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Mathematical Modelling: Theory and Applications 23, Springer, New York, 2009.

[22] D. P. Kouri, *An Approach for the Adaptive Solution of Optimization Problems Governed by Partial Differential Equations with Uncertain Coefficients*, ProQuest LLC, Ann Arbor, MI, 2012, Thesis (Ph.D.)–Rice University. MR 3130817

[23] D. P. Kouri, M. Heinkenschloss, D. Ridzal, and B. G. van Bloemen Waanders, *A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty*, SIAM J. Sci. Comput. **35** (2013), no. 4, A1847–A1879. MR 3073358

[24] D. P. Kouri and T. M. Surowiec, *Risk-averse PDE-constrained optimization using the conditional value-at-risk*, SIAM J. Optim. **26** (2016), no. 1, 365–396. MR 3455142

[25] A. Kunoth and C. Schwab, *Analytic regularity and GPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs*, SIAM J. Control Optim. **51** (2013), no. 3, 2442–2471. MR 3064588

[26] H. J. Kushner and G. G. Yin, *Stochastic approximation algorithms and applications*, Applications of Mathematics (New York), vol. 35, Springer-Verlag, New York, 1997.

[27] G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher, and S. Ulbrich (eds.), *Trends in PDE constrained optimization*, Birkhäuser/Springer, Cham, 2014.

[28] J.L. Lions, *Optimal control of systems governed by partial differential equations*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 1971.

[29] G.J. Lord, C.E. Powell, and T. Shardlow, *An introduction to computational stochastic pdes*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2014.

[30] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, *Robust stochastic approximation approach to stochastic programming*, SIAM Journal on Optimization **19** (2009), no. 4, 1574–1609.

[31] B.T. Polyak and A.B. Juditsky, *Acceleration of stochastic approximation by averaging*, SIAM Journal on Control and Optimization **30** (1992), no. 4, 838–855.

[32] A. Quarteroni, *Numerical models for differential problems*, MS&A, Springer, Milano, 2009.

[33] H. Robbins and S. Monro, *A stochastic approximation method*, Ann. Math. Statist. **22** (1951), no. 3, 400–407.

[34] R. T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions*, J. Bank. Financ. **26** (2002), no. 7, 1443–1471.

[35] E. Rosseel and G. N. Wells, *Optimal control with stochastic PDE constraints and uncertain controls*, Comput. Methods Appl. Mech. Engrg. **213/216** (2012), 152–167. MR 2880511

[36] D. Ruppert, *Efficient estimations from a slowly convergent robbins-monro process*, Tech. report, Cornell University Operations Research and Industrial Engineering, 1988.

[37] C. Schillings, S. Schmidt, and V. Schulz, *Efficient shape optimization for certain and uncertain aerodynamic design*, Computers & Fluids **46** (2011), no. 1, 78 – 87, 10th ICFD Conference Series on Numerical Methods for Fluid Dynamics (ICFD 2010).

[38] M. Schmidt, N. Le Roux, and F. Bach, *Minimizing finite sums with the stochastic average gradient*, Mathematical Programming **162** (2017), no. 1-2, 83–112.

[39] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming*, Society for Industrial and Applied Mathematics, 2009.

[40] H. Tiesler, R. M. Kirby, D. Xiu, and T. Preusser, *Stochastic collocation for optimal control problems with stochastic PDE constraints*, SIAM J. Control Optim. **50** (2012), no. 5, 2659–2682. MR 3022082

[41] A. Van Barel, S. Vandewalle, and P. Robbe, *Robust Optimization of Systems Described by Partial Differential Equations using a Multilevel Monte Carlo Method*, International Multigrid Conference, Bruchsal, Germany, 4-9 December 2016, December 2016.

[42] D. Williams, *Probability with martingales*, Cambridge mathematical textbooks, Cambridge University Press, 1991.

CSQI, Institute of Mathematics, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

E-mail address, M. Martin: matthieu.martin@epfl.ch

E-mail address, S. Krumscheid: sebastian.krumscheid@epfl.ch

E-mail address, F. Nobile: fabio.nobile@epfl.ch

Recent publications:

INSTITUTE of MATHEMATICS
MATHICSE Group

Ecole Polytechnique Fédérale (EPFL)

CH-1015 Lausanne

2018

01.2018 ANA SUSNAJARA, DANIEL KRESSNER:
A fast spectral divide-and-conquer method for banded matrices

02.2018 ELEONORA ARNONE, LAURA AZZIMONTI, FABIO NOBILE, LAURA M. SANGALLI:
Modelling spatially dependent functional data via regression with differential regularization

03.2018 NICCOLO DAL SANTO, SIMONE DEPARIS, ANDREA MANZONI, ALFIO QUARTERONI:
Mutli space reduced basis preconditioners for parametrized Stokes equations

04.2018 MATTHIEU MARTIN, SEBASTIAN KRUMSCHEID, FABIO NOBILE:
Mutli space reduced basis preconditioners for parametrized Stokes equations
Analysis of stochastic gradient methods for PDE-Constrained optimal control problems with uncertain parameters

***