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

M. MARTIN

Criteo AI Lab, 4 Rue des Méridiens, 38130 Échirolles, France

S. KRUMSCHEID

Department of Mathematics, RWTH Aachen, 52062 Aachen, Germany

F. NOBILE

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

Abstract. We consider the numerical approximation of an 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 squared L^2 misfit 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 (OCP) we consider a Finite Element discretization of the underlying PDE, a Monte Carlo sampling method, and gradient-type iterations to obtain the approximate optimal control. We provide full error and complexity analyses of the proposed numerical schemes. In particular we investigate the complexity of a conjugate gradient method applied to the fully discretized OCP (so called Sample Average Approximation), in which the Finite Element discretization and Monte Carlo sample are chosen in advance and kept fixed over the iterations. This is compared with a *Stochastic Gradient* method on a fixed or varying Finite Element discretization, in which the expectation in the computation of the steepest descent direction is approximated by Monte Carlo estimators, independent across iterations, with small sample sizes. We show in particular that the second strategy results in an improved computational complexity. The theoretical error estimates and complexity results are confirmed by numerical experiments.

Keywords. PDE constrained optimization, risk-averse optimal control, optimization under uncertainty, PDE with random coefficients, sample average approximation, stochastic approximation, stochastic gradient, Monte Carlo
AMS subject classifications. 35Q93, 49M99, 65C05, 65N12, 65N30

E-mail addresses: mat.martin@criteo.com, krumscheid@uq.rwth-achen.de, fabio.nobile@epfl.ch.

Date: May 28, 2020.

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, 13, 25, 27, 34]. 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. These uncertainties could appear for instance in the material properties or boundary conditions and are often described probabilistically in terms of random variables or random fields. Such Optimal Control Problems (OCPs) are sometimes also referred to as problems of 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 deterministic volumetric forcing term, so that the controlled solution is as close as possible to a given target function. While there is a vast literature on the numerical approximation of PDE-constrained optimal control problems in the deterministic case (see, e.g., [7, 27] and the references therein), as well as on the numerical approximation of (uncontrolled) PDEs with random coefficients (e.g. [3, 23, 35]), the analysis of PDE constrained control problems under uncertainty is much more recent and incomplete, although the topic has received increasing attention in the last few years.

The different frameworks of PDE-constrained OCPs under uncertainty considered in the literature can be roughly grouped in two categories.

In the first category, the control is *random* [1, 6, 10, 32, 41, 46]. 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. The corresponding optimality system might still be fully coupled in the random parameters, e.g. if the random objective function also involves some statistics of the state variables (e.g. deviation from the nominal response). The dependence on the random parameters is typically approximated either by polynomial chaos expansions or sampling-based Monte Carlo (MC) techniques. This is, for example, considered in [32] 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. In [10] the authors combine a stochastic collocation method with a Finite Element (FE) based reduced basis method to alleviate the computational effort. The works [6, 41, 46] 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. Sampling-based 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, 22, 28, 29, 30, 47]. 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. The precise notion of a risk is problem dependent and thus has to be modeled appropriately. In this context, risk typically refers to a suitable statistical measure of the objective function to be minimized, such as those involving expectations, expectations plus variances, a quantile, or a conditional expectation above a quantile (so called Conditional Value at Risk, CVaR [40]). The corresponding OCP often leads to a fully coupled optimality system of equations in the random parameters. It is noteworthy that the idea of minimizing a risk to obtain a solution with favorable properties goes back to the origins of robust optimization [45].

Numerical methods for OCPs of the second category typically depend on the choice of the risk measure. For example, in [2] the authors consider a risk measure based on the mean and variance of a scalar objective function and they use second order Taylor expansions combined with randomized estimators to reduce the computational effort. The work [47] contains a study of a risk measure that involves the expected squared L^2 misfit between the state and a target function, with an additional penalty term involving the squared L^2 deviation of the state form its mean value. The authors propose 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 similar risk measure and propose a reduced basis method on the space of controls to significantly reduce the computational effort. A more general class of risk measures (including the CVaR) for OCPs has been considered in [31], where also the corresponding optimality system of PDEs are derived. The subsequent work [29] introduces a trust-region (Newton) conjugate gradient algorithm combined with an adaptive sparse grid collocation as PDE discretization in the stochastic space for the numerical treatment of these more general OCPs. For the robust OCP with the CVaR as risk measure the study [30] introduces derivative-based optimization methods that are build upon introducing smooth approximations to the CVaR. Finally, in [22] 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 and consider a similar risk averse OCP as in [9, 47] which consists in minimizing the expected squared L^2 misfit between the state and a given target function as objective function, additionally equipped with an L^2 regularization on the (deterministic) control. For this setting we consider numerical gradient based methods, either deterministic or stochastic, where adjoint calculus is used to represent the gradient of the objective function. Both the primal problem and the adjoint problem are discretized by finite elements and Monte Carlo estimators are used to approximate expectations defining the risk measure. The reason for studying sampling-based Monte Carlo approximations instead of polynomial chaos type methods 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 and complexity analysis for the considered gradient based methods, accounting for the three sources of errors, namely, the Finite Element approximation, the statistical Monte Carlo error, and the error due to the finite number of gradient based iterations.

We note that other error analyses have been presented before, such as [10] for the case of a random control with a discretization in the physical space by Finite Elements and in probability by a stochastic collocation, as well as [22] for the case of a deterministic boundary control that minimizes a quadratic risk measure, using a Finite Element discretization both in physical space and in probability. Finally we mention the recent work [24] where the authors consider as a risk measure the same expected quadratic loss function as in this work and study a quasi-Monte Carlo approximation (i.e., a deterministic quadrature in the probability space) of the expected loss which may offer a further complexity improvement, provided that the system's state equation is sufficient regular with respect to the uncertain parameters. In contrast, here we focus on the subtle interplay of Finite Element discretization errors, Monte Carlo sampling errors and different numerical optimization techniques in the context of general-purpose methods.

The first method that we consider is a gradient based method (e.g. conjugate gradient) on a fully discretized version of the OCP (so called Sample Average Approximation – SAA), in which the Finite Element discretization and the Monte Carlo sample are chosen initially and kept fixed over the iterations. If N is the

sample size of the Monte Carlo estimator, this method entails the solution of N state and N adjoint 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 gradient methods in which the gradient is re-sampled independently at each iteration and the Finite Element mesh size can be refined along the iterations. At each iteration this corresponds to taking an independent Monte Carlo estimator with only one realization ($N = 1$) or a very small, fixed sample size ($N = \bar{N}$) independently of the required tolerance, with possibly a finer Finite Element mesh. We follow, in particular, the Robbins–Monro strategy [37, 39, 42] of reducing progressively the step-size to achieve convergence of the Stochastic Gradient iterations. These *Stochastic Gradient* (SG) techniques have been extensively applied to machine learning problems [14, 16, 19, 33], but have not yet been used much for PDE-constrained optimization under uncertainty. Here, we show that a Stochastic Gradient method improves the complexity of the conjugate gradient (or equivalent) method applied to the fully discretized OCP by a logarithmic factor. Although the computational gain is not dramatic, we see potential in this approach as only one state problem and one adjoint 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, but mention the related recent work [20] on mesh refinement approaches in the context of a stochastic gradient method for PDE-constrained OCPs subject to uncertainties.

The rest of the paper is organized as follows. In Section 2 we introduce the optimal control problem under uncertainty and recall its well posedness as well as the corresponding optimality conditions. In Sections 3, 4, and 5 we then introduce the Finite Element discretization, the Monte Carlo approximation, and the gradient based methods, respectively, including their full error analysis. In particular, Theorem 2 in Section 5 provides an error bound for the conjugate gradient (or equivalent) method applied to the fully discrete OCP, whereas Corollary 1 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 3 and the corresponding complexity result in Corollary 2), whereas in Section 7 we analyze the Stochastic Gradient version in which the Finite Element mesh is refined over the iterations (the main result being stated in Theorem 5 and Corollary 3). 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 state problem that will be part of the OCP discussed in the following. Specifically, we consider the problem of finding the solution $y\colon D \times \Gamma \to \mathbb{R}$ 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 possibly stochastic source term (which could include a deterministic control term).

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 $L^p(D)$, $1 \leq p \leq +\infty$, denote the space of p-Lebesgue integrable functions. Throughout this work, we will denote by $\|\cdot\| \equiv \|\cdot\|_{L^2(D)}$ the $L^2(D)$-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

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

and

$$H_0^1(D) = \{y \in H^1(D), \quad y|_{\partial D} = 0\},$$

on which a Poincaré inequality holds: $\exists C_p > 0 : \ \|y\| \leq C_p \|\nabla y\|, \forall y \in H_0^1(D)$. We use the equivalent H^1-norm on the space $H_0^1(D)$ defined by $\|y\|_{H_0^1(D)} = \|\nabla y\|$ for any $y \in H_0^1(D)$, and we denote by $H^{-1}(D) = \left(H_0^1(D)\right)^*$ the topological dual of $H_0^1(D)$. For $r \in \mathbb{N}$ we further recall the subspace $H^r(D)$ of $L^2(D)$ composed of functions with all weak 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

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

respectively, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is a multi-index. 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

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

equipped with the norm $\|y\|_{L^p(\Gamma,\mathcal{V})} = \left(\int_\Gamma \|y(\omega)\|_{\mathcal{V}}^p \mathrm{d}P(\omega)\right)^{\frac{1}{p}}$; see, e.g., [18] 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, 35].

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 $y \in L^2(\Gamma, H_0^1(D))$ s.t.*

$$\|y(\cdot, \omega)\|_{H_0^1(D)} \leq \frac{1}{a_{\min}} \|\phi(\cdot, \omega)\|_{H^{-1}(D)} \quad \text{for a.e. } \omega \in \Gamma$$

$$\text{and } \|y\|_{L^2(\Gamma, H_0^1(D))} \leq \frac{1}{a_{\min}} \|\phi\|_{L^2(\Gamma, H^{-1}(D))}.$$

Finally, as we will occasionally need H^2-regularity in the following Sections, we also introduce the following sufficient conditions 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 is such that $a(\cdot, \omega) \in C^{0,1}(D)$ for a.e. $\omega \in \Gamma$ with $\operatorname{ess\,sup}_\omega \|a(\cdot, \omega)\|_{C^{0,1}(D)} < \infty$.*

Then, using standard regularity arguments for elliptic PDEs, one can prove the following result [18, 21].

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 state problem for the OCP as the elliptic PDE (1), by specializing its right hand side to:

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

with source term $g \in L^2(D)$ and control $u \in L^2(D)$. Hereafter, we use the notation $U = L^2(D)$ to denote the set of all admissible (deterministic) control functions, and $Y = H_0^1(D)$ to denote the state space of solutions to (2). To emphasize the dependence of the solution of 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 functional

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

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 $\beta > 0$ 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 *unconstrained* OCP, of determining the optimal control $u_\star$ so that

$$u_\star \in \arg\min_{u \in U} J(u), \quad \text{s.t.} \quad y_\omega(u) \in Y \quad \text{solves} \quad (2) \quad \text{a.s.} \tag{4}$$

As we aim at minimizing the objective 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 $u_\star$ satisfies

$$\langle \nabla J(u_\star), v \rangle = 0 \quad \forall v \in U. \tag{5}$$

Here, $\nabla J(u)$ denotes the $L^2(D)$-functional representation of the Gateaux derivative of J at $u \in U$, namely

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

Existence and uniqueness results for the OCP (4) can be obtained as a particular case of the more general results in, e.g., the work [31]. We state the result in the next Lemma, without proof.

Lemma 3. *Suppose Assumption 1 holds. Then the OCP (4) admits a unique control $u_\star \in U$. Moreover*

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

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

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

One can derive a similar expression as (6) for the gradient of the functional f for a.e. $\omega \in \Gamma$, namely: $\nabla f(u,\omega) = \beta u + p_\omega(u)$. Consequently, in the setting of problem (4), we have

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

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

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

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

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

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

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

3. Finite Element approximation in physical space

In this section we introduce the semi-discrete OCP obtained by approximating the underlying PDE by a Finite Element (FE) method and recall an a-priori error bound on 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. $Y^h = \{y \in C^0(\overline{D}) : y|_K \in \mathbb{P}_r(K) \quad \forall K \in \tau_h, y|_{\partial D} = 0\} \subset Y = H_0^1(D)$. Finally, we set $U^h = Y^h$. We can then reformulate (10) as a finite dimensional OCP in the FE space:

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

Remark 1. *The choice $U^h = Y^h$ is natural for this problem. In fact, one could consider the OCP in which the PDE is discretized in Y^h, whereas the control $u \in U$ is not discretized. It is not difficult to show that the optimal control for such OCP is actually finite dimensional and belongs to Y^h, thus leading to the equivalent formulation* (11).

For the discrete OCP (11) we have analogous well-posedness and optimality results as those stated in Lemma 3 for the continuous problem.

Lemma 4. *The discrete OCP* (11) *admits a unique solution $u^h_\star \in U^h$ and ∇J^h can be characterized as*

$$\nabla J^h(u^h) = \beta u^h + \mathbb{E}[p^h_\omega(u^h)] \tag{12}$$

where $p^h_\omega(u^h)$ is the solution of the FE adjoint problem

$$\text{find } p^h_\omega(u^h) \in Y^h \text{ s.t. } b_\omega(v^h, p^h_\omega(u^h)) = \langle y^h_\omega(u^h) - z_d, v^h\rangle \quad \forall v^h \in Y^h.$$

Notice, in particular, that, since $U^h = Y^h$, it follows that $\mathbb{E}[p^h(u^h)] \in U^h$ for any $u^h \in U^h$, hence $\nabla J^h(u^h) = \beta u^h + \mathbb{E}[p^h(u^h)] \in U^h$.

Following similar arguments as in [27, Thms. 3.4, 3.5] and using the optimality condition and the weak form of the state and adjoint problems, we can prove the following.

Lemma 5. *Let $u_\star$ be the optimal control solution of problem* (10) *and denote by $u_\star^h$ the solution of the discrete OCP* (11)*. Then it holds that*

(13)
$$\frac{\beta}{2}\|u_\star - u_\star^h\|^2 + \frac{1}{2}\mathbb{E}[\|y(u_\star) - y^h(u_\star^h)\|^2] \leq \frac{1}{2\beta}\mathbb{E}[\|p(u_\star) - \widetilde{p}^h(u_\star)\|^2] + \frac{1}{2}\mathbb{E}[\|y(u_\star) - y^h(u_\star)\|^2],$$

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

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

Moreover, there exists a constant $C > 0$ independent of h such that

$$\begin{aligned}&\|u_\star - u_\star^h\|^2 + \mathbb{E}[\|y(u_\star) - y^h(u_\star^h)\|^2] + h^2\mathbb{E}[\|y(u_\star) - y^h(u_\star^h)\|^2_{H_0^1}] \\ &\leq C\{\mathbb{E}[\|p(u_\star) - \widetilde{p}^h(u_\star)\|^2] + \mathbb{E}[\|y(u_\star) - y^h(u_\star)\|^2] + h^2\mathbb{E}[\|y(u_\star) - y^h(u_\star)\|^2_{H_0^1}]\}.\end{aligned} \tag{15}$$

Proof. The result in the deterministic case is detailed in [27, Thms. 3.4, 3.5]. We can thus write the inequalities (13) and (14) for almost every realization $\omega \in \Gamma$, and then take the expectation to conclude. □

The FE error $\|u_\star - u_\star^h\|$ on the optimal control is thus completely determined by the FE approximation properties of the state and adjoint problems. The next result also follows by standard arguments (see e.g. [27]) and shows that, for smooth data, the L^2 error $\|u_\star - u_\star^h\|$ on the optimal control converges at rate $O(h^{r+1})$.

Lemma 6. *Let Assumptions 1-2 hold and suppose that $y(u_\star), p(u_\star) \in L^2(\Gamma, H^{r+1}(D))$. Then we have*

$$\begin{aligned}&\|u_\star - u_\star^h\|^2 + \mathbb{E}[\|y(u_\star) - y^h(u_\star^h)\|^2] + h^2\mathbb{E}[\|y(u_\star) - y^h(u_\star^h)\|^2_{H_0^1}] \\ &\qquad \leq Ch^{2r+2}\{\mathbb{E}[|y_\omega(u_\star)|^2_{H^{r+1}}] + \mathbb{E}[|p_\omega(u_\star)|^2_{H^{r+1}}]\}.\end{aligned} \tag{16}$$

In view of the analysis that will be presented later, we state a Lipschitz and a strong convexity result for the functional $f(u, \omega)$ for a.e. $\omega \in \Gamma$, as well as its discrete version $f^h(u^h, \omega) := \frac{1}{2}\|y^h_\omega(u^h) - z_d\|^2 + \frac{\beta}{2}\|u^h\|^2$. Proofs of these results can be found in [36].

Lemma 7 (Lipschitz condition)**.** *For the elliptic problem* (2) *and $f(u, \omega)$ as in* (3) *it holds that:*

$$\|\nabla f(u_1, \omega) - \nabla f(u_2, \omega)\| \leq L\|u_1 - u_2\| \quad \forall u_1, u_2 \in U \text{ and a.e. } \omega \in \Gamma, \tag{17}$$

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

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

Lemma 8 (Strong Convexity)**.** *For the elliptic problem* (2) *and $f(u, \omega)$ as in* (3) *it holds that*

$$\frac{l}{2}\|u_1 - u_2\|^2 \leq \langle \nabla f(u_1, \omega) - \nabla f(u_2, \omega), u_1 - u_2\rangle \quad \forall u_1, u_2 \in U \text{ and a.e. } \omega \in \Gamma, \tag{18}$$

with $l = 2\beta$. The same estimate holds for the corresponding Finite Element approximation that yields (11)*, namely:*

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

4. Approximation in probability

In this section we consider the semi-discrete (i.e., approximation in probability only) optimal control problem obtained by replacing the exact expectation $\mathbb{E}[\cdot]$ in (3) by a suitable quadrature formula $\widehat{E}[\cdot]$. Specifically, for a random variable $X : \Gamma \to \mathbb{R}$, $\omega \mapsto X(\omega)$, let $\widehat{E}[X] = \sum_{i=1}^{N} \zeta_i X(\omega_i)$ be a quadrature formula, where ζ_i denote the quadrature weights and $\omega_i \in \Gamma$ the quadrature points. The semi-discrete problem then reads:

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

The quadrature formula $\widehat{E}[\cdot]$ could either be based on deterministic quadrature points or randomly distributed points leading, in this case, to a Monte Carlo type approximation. In the following, we detail the case of a Monte Carlo type quadrature, whereas the case of a deterministic Gaussian-type quadrature is addressed in Appendix A. It is noteworthy that, although we present the results only for the semi-discrete problem (i.e., continuous in space, discrete in probability) for the sake of notation, they extend straightforwardly to the fully discrete problem in probability and physical space. Indeed, the fully discrete problem is obtained by replacing the (spatial) functions and corresponding functions spaces in (19) by their finite dimensional Finite Element approximations.

In the case of a Monte Carlo (MC) approximation, the quadrature formula reads $E^{\overrightarrow{\omega}}_{MC}[X] := \frac{1}{N}\sum_{i=1}^{N} X(\omega_i)$, where $\overrightarrow{\omega} = \{\omega_i\}_{i=1}^{N}$ is a collection of independent and identically distributed (i.i.d.) points drawn randomly on Γ according to the probability measure P. We recall that the use of MC type approximations might be advantageous over a quadrature/collocation approach in cases where the state and adjoint solutions are rough or highly oscillatory, which is, for example, the case when $a(\cdot,\cdot)$ is a rough random field and/or has a short correlation length. Moreover, the Monte Carlo quadrature formula has always *positive* weights, which is an important feature to guarantee that the approximate functional $\widehat{J}$ preserves the strong convexity property. We stress that, when using a Monte Carlo quadrature formula, the optimal control $\widehat{u}_\star$ is a stochastic function since it depends on the N i.i.d. random points $\overrightarrow{\omega} = \{\omega_i\}_{i=1}^{N}$. The next theorem gives an error bound on the approximate optimal control of the OCP (19).

Theorem 1. *Let $\widehat{u}_\star$ be the optimal control of problem* (19) *with $\widehat{E} = E^{\overrightarrow{\omega}}_{MC}$ and $u_\star$ be the exact optimal control of the continuous problem* (10)*. Then we have*

$$\frac{\beta}{2}\mathbb{E}[\|\widehat{u}_\star - u_\star\|^2] + \mathbb{E}[\|y(u_\star) - y(\widehat{u}_\star)\|^2] \le \frac{1}{N}\frac{1}{2\beta}\mathbb{E}[\|p(\widehat{u}_\star)\|^2].$$

Proof. The two optimality conditions for OCPs (10) and (19) read

$$\langle \nabla J(u_\star), v\rangle = 0, \quad \text{and} \quad \langle \nabla J_{MC}(\widehat{u}_\star), v\rangle = 0 \quad \forall v \in U \tag{20}$$

respectively, where $\nabla J_{MC}(\widehat{u}_\star) = \beta\widehat{u}_\star + E^{\overrightarrow{\omega}}_{MC}[p(\widehat{u}_\star)]$ and $p(\widehat{u}_\star) = p_\omega(\widehat{u}_\star)$ denotes the solution of the adjoint problem

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

Choosing $v = \widehat{u}_\star - u_\star$ in (20) and subtracting the two optimality conditions, we obtain:

$$\langle \beta(u_\star - \widehat{u}_\star) + \mathbb{E}[p(u_\star)] - E^{\overrightarrow{\omega}}_{MC}[p(\widehat{u}_\star)], \widehat{u}_\star - u_\star\rangle = 0,$$

which implies

$$\beta\|u_\star - \widehat{u}_\star\|^2 = \langle \mathbb{E}[p(u_\star)] - E^{\vec{\omega}}_{MC}[p(u_\star)], \widehat{u}_\star - u_\star\rangle + \langle E^{\vec{\omega}}_{MC}[p(u_\star)] - E^{\vec{\omega}}_{MC}[p(\widehat{u}_\star)], \widehat{u}_\star - u_\star\rangle. \tag{21}$$

The first term on the right hand side of (21) can be bounded as

$$\langle \mathbb{E}[p(u_\star)] - E^{\vec{\omega}}_{MC}[p(u_\star)], \widehat{u}_\star - u_\star\rangle \le \frac{1}{2\beta}\|\mathbb{E}[p(u_\star)] - E^{\vec{\omega}}_{MC}[p(u_\star)]\|^2 + \frac{\beta}{2}\|\widehat{u}_\star - u_\star\|^2 .$$

To bound the second term, we first notice that for any $i = 1, \dots, N$

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

which leads to

$$\langle \widehat{u}_\star - u_\star, E^{\vec{\omega}}_{MC}[p(u_\star)] - E^{\vec{\omega}}_{MC}[p(\widehat{u}_\star)]\rangle = -E^{\vec{\omega}}_{MC}[\|y(u_\star) - y(\widehat{u}_\star)\|^2] .$$

Finally we take the expectation of (21) with respect to (w.r.t.) the random sample $\vec{\omega} = \{\omega_i\}_{i=1}^N$ and exploit the fact that the Monte Carlo estimator is unbiased, that is $\mathbb{E}[E^{\vec{\omega}}_{MC}[X]] = \mathbb{E}[X]$ for any random variable $X \in L^1(\Gamma)$. We thus find

$$\begin{aligned}\frac{\beta}{2}\mathbb{E}[\|\widehat{u}_\star - u_\star\|^2 + \mathbb{E}[\|y(u_\star) - y(\widehat{u}_\star)\|^2] &\le \frac{1}{2\beta}\mathbb{E}[\|\mathbb{E}[p(\widehat{u}_\star)] - E^{\vec{\omega}}_{MC}[p(\widehat{u}_\star)]\|^2] \\ &= \frac{1}{2\beta}\mathbb{E}[\|\frac{1}{N}\sum_{i=1}^N p_{\omega_i}(\widehat{u}_\star) - \mathbb{E}[p(\widehat{u}_\star)]\|^2] \\ &= \frac{1}{2\beta}\mathbb{E}[\frac{1}{N^2}\sum_{i=1}^N \|p_{\omega_i}(\widehat{u}_\star) - \mathbb{E}[p(\widehat{u}_\star)]\|^2] \\ &= \frac{1}{2\beta}\frac{1}{N}\mathbb{E}[\|p(\widehat{u}_\star) - \mathbb{E}[p(\widehat{u}_\star)]\|^2] \\ &\le \frac{1}{2\beta}\frac{1}{N}\mathbb{E}[\|p(\widehat{u}_\star)\|^2] ,\end{aligned}$$

which concludes the proof. □

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

5. Numerical solution of the fully discrete problem

Now we focus on a class of optimization methods to approximate the fully discrete minimization problem obtained by using the Monte Carlo estimator to approximate the expectation in (11) and a FE approximation of the state and adjoint equations, as discussed in the previous two sections. That is, here we consider the fully discrete OCP:

$$\begin{cases} \min_{u^h \in U^h} J_{MC}(u^h) = \frac{1}{2}E^{\vec{\omega}}_{MC}[\|y^h_\omega(u^h) - z_d\|^2] + \frac{\beta}{2}\|u^h\|^2 \\ \text{s.t.} \quad y^h_{\omega_i}(u^h) \in Y^h \quad \text{and} \\ b_{\omega_i}(y^h_{\omega_i}(u^h), v^h) = \langle g + u^h, v^h\rangle \quad \forall v^h \in Y^h, \quad i = 1, \dots, N. \end{cases} \tag{22}$$

The N constraints in (22) can be written in algebraic form as

$$A_i \mathbf{y}_i = \mathbf{g} + M\mathbf{u} ,$$

where $\mathbf{u} \in \mathbb{R}^{N_h}$ is the vector of the N_h degrees of freedom corresponding to the control $u^h \in U^h$, $\mathbf{y}_i \in \mathbb{R}^{N_h}$ is the vector of degrees of freedom corresponding to the finite element state solution $y^h_{\omega_i}(u^h) \in Y^h$, $M \in \mathbb{R}^{N_h \times N_h}$ is the FE mass matrix, and

$A_i \in \mathbb{R}^{N_h \times N_h}$ is the FE stiffness matrix corresponding to the diffusion coefficient $a(\cdot, \omega_i)$. Defining the block matrices and vectors

$$\mathcal{A} = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_N \end{bmatrix}, \qquad \mathcal{M} = \begin{bmatrix} M & & \\ & \ddots & \\ & & M \end{bmatrix}, \qquad \mathbb{1} = \begin{bmatrix} I_{N_h} \\ \vdots \\ I_{N_h}, \end{bmatrix}$$

$$\vec{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix}, \qquad \vec{p} = \begin{bmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_N \end{bmatrix}, \qquad \vec{g} = \mathbb{1}\mathbf{g}, \qquad \vec{z}_d = \mathbb{1}\mathbf{z}_d,$$

where the $\mathbf{p}_i$ solve the adjoint systems $A_i^T \mathbf{p}_i = M\mathbf{y}_i - \mathbf{z}_d$, the optimality condition for (22) reads $\beta M\mathbf{u} + \frac{1}{N}\sum_{i=1}^N M\mathbf{p}_i = 0$, which leads to the coupled linear system

$$\left[\begin{array}{c|c|c} \mathcal{A} & 0 & -\mathcal{M}\mathbb{1} \\ \hline -\mathcal{M} & \mathcal{A}^T & 0 \\ \hline 0 & \mathbb{1}^T\mathcal{M} & N\beta M \end{array}\right] \left[\begin{array}{c} \vec{y} \\ \hline \vec{p} \\ \hline \mathbf{u} \end{array}\right] = \left[\begin{array}{c} \vec{g} \\ \hline -\vec{z}_d \\ \hline 0 \end{array}\right]. \tag{23}$$

By eliminating the state and adjoint unknowns, this can be recast into a linear system in the control variable only

$$G\mathbf{u} = \boldsymbol{\chi}\,, \tag{24}$$

with $G = N\beta M + \mathbb{1}^T\mathcal{M}\mathcal{A}^{-T}\mathcal{M}\mathcal{A}^{-1}\mathcal{M}\mathbb{1}$ and $\boldsymbol{\chi} = \mathbb{1}^T\mathcal{M}\mathcal{A}^{-T}\left(\vec{z}_d - \mathcal{M}\mathcal{A}^{-1}\vec{g}\right)$.

System (24) can, for example, be solved by an iterative method such as gradient or conjugate gradient type methods. At each iteration, a matrix-vector multiplication will involve the solution of N state and N adjoint equations, resulting in $2N$ PDE solves per iteration. Both iterative solvers are examples of solution schemes that offer an exponential convergence in the number of iterations. Alternative to the formulation (24) one could rewrite system (23) in the form

$$\left[\begin{array}{c|c} C & B^T \\ \hline B & 0 \end{array}\right] \left[\begin{array}{c} \vec{w} \\ \hline \vec{p} \end{array}\right] = \left[\begin{array}{c} \vec{v} \\ \hline \vec{g} \end{array}\right], \qquad \text{with } \vec{w} = \begin{bmatrix} \vec{y} \\ \mathbf{u} \end{bmatrix}, \tag{25}$$

which could be solved, e.g. by GMRES or MINRES iterative methods. In the next Theorem 2 we analyze the complexity of the fully discrete problem in terms of computational work needed to achieve a given tolerance. Instead of particularizing the result to one specific iterative solver, we make the assumption that an exponentially convergent iterative solver is employed. More precisely:

Assumption 3. *Let $\widehat{u}_\star^h$ be the exact solution to* (22)*. Furthermore, let $\widehat{u}_j^h$ denote the approximate solution to the OCP* (22) *obtained after j iterations of the iterative solver used to solve* (22)*. We assume that the chosen iterative solver satisfies*

$$\mathbb{E}[\|\widehat{u}_j^h - \widehat{u}_\star^h\|^2] \leq C_1 e^{-\rho j}, \quad \forall j \in \mathbb{N}\,, \tag{26}$$

for some constants $C_1, \rho > 0$ that are independent of h and N.

This assumption is sound since the condition number of the matrix G in (24) can be bounded uniformly in h and N and scales as β^{-1}. Similarly, the system (25) can be optimally preconditioned, so that the exponential convergence rate does not depend on the discretization parameters.

Based on Assumption 3 concerning the iterative solver, we now provide an error bound for the approximate solution $\widehat{u}_j^h$, as a function of all discretization parameters j, h, and N.

Theorem 2. *Let $u_\star$ be the solution of the optimal control problem* (10)*. Moreover, let $\widehat{u}_j^h$ be the j-th iteration of a linear solver applied to* (24) *and suppose that the solver satisfies Assumption 3. Then under the assumptions of Lemma 6, there exist constants $C_1, C_2, C_3 > 0$ independent of h and N such that*

$$\mathbb{E}[\|\widehat{u}_j^h - u_\star\|^2] \leq C_1 e^{-\rho j} + \frac{C_2}{N} + C_3 h^{2r+2} \,. \tag{27}$$

Proof. The global error can be decomposed as follows:

$$\mathbb{E}[\|\widehat{u}_j^h - u_\star\|^2] \leq 3\underbrace{\mathbb{E}[\|\widehat{u}_j^h - \widehat{u}_\star^h\|^2]}_{\text{linear solver}} + 3\underbrace{\mathbb{E}[\|\widehat{u}_\star^h - u_\star^h\|^2]}_{\text{MC}} + 3\underbrace{\mathbb{E}[\|u_\star^h - u_\star\|^2]}_{\text{FE error}}.$$

The first term $\mathbb{E}[\|\widehat{u}_j^h - \widehat{u}_\star^h\|^2]$ quantifies the convergence of the finite dimensional optimization algorithm, which is exponential w.r.t. the iteration number thanks to the hypotheses.

The second term $\mathbb{E}[\|\widehat{u}_\star^h - u_\star^h\|^2]$ accounts for the standard MC error and can be controlled as in Theorem 1 (applied to the FE approximation) leading to

$$\mathbb{E}[\|\widehat{u}_\star^h - u_\star^h\|^2] \leq \frac{1}{\beta^2 N}\mathbb{E}[\|p(\widehat{u}_\star^h)\|^2].$$

Finally, the term $\mathbb{E}[\|u_\star^h - u_\star\|^2]$ can be controlled by the result in Lemma 6, namely by

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

so that the claim follows. □

We conclude this Section by analyzing the computational complexity of solving the fully discrete OCP (22), or equivalently (24), using an exponentially convergent iterative solver. We assume that the state and adjoint problems, using a triangulation with mesh size h, can be solved 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 1. *In order to achieve a given tolerance $O(tol)$, i.e. to guarantee that $\mathbb{E}[\|\widehat{u}_j^h - u_\star\|^2] \lesssim tol^2$, the total required computational work W 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 (27). This leads to the choices:

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

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. □

6. Stochastic Gradient with fixed mesh size.

In the previous Section, we have considered an approach to approximately solve the OCP (10) in which the Monte Carlo sample size N is fixed a-priori, based on accuracy requirements, the sample is generate once and for all, and then the coupled system (23) is solved by an iterative scheme, each iteration involving the solution of N primal and N adjoint problems.

As an alternative, in this section we consider an approach based on stochastic optimization ideas. Specifically, we will use randomized methods known in the literature as Stochastic Approximation (SA) or Stochastic Gradient (SG) [16, 38, 39, 44, 45]. At the basis of these methods is a steepest decent algorithm to tackle

the optimization problem. For example, 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 a single evaluation $\nabla f(\cdot, \omega_j)$, where the random variable ω_j is re-sampled independently at each iteration of the steepest-descent method:

$$u_{j+1} = u_j - \tau_j \nabla f(u_j, \omega_j). \tag{28}$$

Here, τ_j is the step-size of the algorithm (also called *learning rate*) and decreases as $1/j$, over the iterations, in the usual approach.

Alternatively, the single evaluation $\nabla f(\cdot, \omega_j)$ can be replaced by a sample average over N_j i.i.d. realizations (so called mini-batches [12, 15, 17]) at every iteration, which are drawn independently of the previous iterations. More precisely, let $\overrightarrow{\omega_j} = (\omega_j^{(1)}, \cdots, \omega_j^{(N_j)}) \sim P^{\otimes N_j}$, then we define the recursion as

$$u_{j+1} = u_j - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \cdot)], \tag{29}$$

where $E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u, \cdot)] = \frac{1}{N_j} \sum_{i=1}^{N_j} \nabla f(u, \omega_j^{(i)})$ is the usual Monte Carlo estimator using a sample of size N_j at iteration j. 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. We first analyze the convergence of the Stochastic Gradient algorithm (29) when applied to the OPC (10) in the continuous setting, i.e. with no Finite Element discretization. The proof of the next theorem follows closely the general one in [45, Sect. 5.9], although here we do not assume uniform boundedness of $\mathbb{E}[\|\nabla f(u, \cdot)\|^2]$ with respect to u, nor do we project the control u onto a bounded set at each iteration, which leads to slight technical modifications. For the sake of completeness, we give the full proof of the theorem.

Theorem 3. *Let $u_\star$ be the solution of the continuous OCP* (10) *and denote by u_j the j-th iterate of* (29)*. Then it holds that*

$$\mathbb{E}[\|u_{j+1} - u_\star\|^2] \leq c_j \mathbb{E}[\|u_j - u_\star\|^2] + \frac{2\tau_j^2}{N_j} \mathbb{E}[\|\nabla f(u_\star, \omega)\|^2], \tag{30}$$

with $c_j := 1 - \tau_j l + L^2 \Big(1 + \frac{2}{N_j}\Big) \tau_j^2$ and L, l the convexity and Lipschitz constants defined in Lemmas 8 and 7, respectively.

Proof. Using inequalities (17) and (18), we can derive a recursive formula to control the error between successive iterations. Let us introduce the filtration $\mathcal{F}_j = \sigma\{\overrightarrow{\omega_k},\ k \leq j-1\}$, i.e. the sigma algebra generated by all random variables used up to iteration $j-1$. Notice that u_j is measurable with respect to $\mathcal{F}_j$. Moreover, we introduce the notation $\mathbb{E}[\cdot|\mathcal{F}_j]$ to denote conditional expectation with respect to such filtration. Using the fact that $\mathbb{E}[\nabla f(u_\star, \cdot)] = 0$, we have:

$$\begin{aligned}
u_{j+1} - u_\star &= u_j - u_\star - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \cdot)] + \tau_j \mathbb{E}[\nabla f(u_\star, \cdot)] \\
&= u_j - u_\star - \tau_j \mathbb{E}[\nabla f(u_j, \cdot)|\mathcal{F}_j] + \tau_j \mathbb{E}[\nabla f(u_\star, \cdot)] + \tau_j \big(\mathbb{E}[\nabla f(u_j, \cdot)|\mathcal{F}_j] - E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \cdot)]\big) \\
&= u_j - u_\star - \tau_j T_1 - \tau_j T_2,
\end{aligned}$$

with $T_1 := \mathbb{E}[\nabla f(u_j, \cdot)|\mathcal{F}_j] - \mathbb{E}[\nabla f(u_\star, \cdot)]$ and $T_2 := E_{MC}^{\overrightarrow{\omega_j}}[\nabla f(u_j, \cdot)] - \mathbb{E}[\nabla f(u_j, \cdot)|\mathcal{F}_j]$. Hence,

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

Concerning the terms containing T_1, we find:

$$\begin{aligned}
\|T_1\|^2 &= \|\mathbb{E}\left[\nabla f(u_j,\cdot)|\mathcal{F}_j\right] - \mathbb{E}[\nabla f(u_\star,\cdot)]\|^2 \\
&= \|\mathbb{E}\left[\nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)|\mathcal{F}_j\right]\|^2 \\
&= \int_D \left(\mathbb{E}\left[\nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)|\mathcal{F}_j\right]\right)^2 \mathrm{d}x \\
&\leq \int_D \mathbb{E}\left[|\nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)|^2|\mathcal{F}_j\right] \mathrm{d}x \\
&= \mathbb{E}\left[\|\nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)\|^2|\mathcal{F}_j\right] \\
&\leq L^2\mathbb{E}\left[\|u_j - u_\star\|^2|\mathcal{F}_j\right],
\end{aligned}$$

where we have used Jensen's inequality for conditional expectation: $\phi(\mathbb{E}\left[X|\mathcal{F}_j\right]) \leq \mathbb{E}\left[\phi(X)|\mathcal{F}_j\right]$ for ϕ convex; see, e.g., [48]. Taking now the full expectation yields:

$$\mathbb{E}[\|T_1\|^2] \leq L^2\mathbb{E}\left[\mathbb{E}\left[\|u_j - u_\star\|^2|\mathcal{F}_j\right]\right] = L^2\mathbb{E}[\|u_j - u_\star\|^2],$$

and

$$\begin{aligned}
\mathbb{E}[\langle u_j - u_\star, T_1\rangle] &= \mathbb{E}[\langle u_j - u_\star, \mathbb{E}\left[\nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)|\mathcal{F}_j\right]\rangle] \\
&= \mathbb{E}\left[\mathbb{E}\left[\langle u_j - u_\star, \nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)\rangle|\mathcal{F}_j\right]\right] \\
&\geq \mathbb{E}\left[\mathbb{E}\left[\frac{l}{2}\|u_j - u_\star\|^2|\mathcal{F}_j\right]\right] \qquad \text{[Strong Convexity (18)]} \\
&= \frac{l}{2}\mathbb{E}[\|u_j - u_\star\|^2].
\end{aligned}$$

We now focus on the term T_2 and notice that it can be written as

$$T_2 = \frac{1}{N_j}\sum_{i=1}^{N_j}\left(\nabla f(u_j,\omega_j^{(i)}) - \mathbb{E}\left[\nabla f(u_j,\cdot)|\mathcal{F}_j\right]\right) = \frac{1}{N_j}\sum_{i=1}^{N_j} Y_i,$$

with $Y_i = \nabla f(u_j,\omega_j^{(i)}) - \mathbb{E}\left[\nabla f(u_j,\cdot)|\mathcal{F}_j\right]$. We have then

$$\mathbb{E}\left[\|T_2\|^2\right] = \mathbb{E}\left[\int_D\left(\frac{1}{N_j}\sum_{i=1}^{N_j} Y_i\right)^2\right] = \mathbb{E}\left[\int_D \frac{1}{N_j^2}\sum_{i,l=1}^{N_j} Y_iY_l\right] = \mathbb{E}\left[\int_D \frac{1}{N_j^2}\sum_{i,l=1}^{N_j}\mathbb{E}\left[Y_iY_l|\mathcal{F}_j\right]\right].$$

Observe that, conditional upon $\mathcal{F}_j$, the random variables Y_i, $i = 1,\ldots,N_j$, are mutually independent and have zero mean, i.e. $\mathbb{E}\left[Y_i|\mathcal{F}_j\right] = 0$ and $\mathbb{E}\left[Y_iY_j|\mathcal{F}_j\right] = 0$ when $i \neq j$. Therefore it follows that

$$\begin{aligned}
\mathbb{E}\left[\|T_2\|^2\right] &= \mathbb{E}\left[\int_D \frac{1}{N_j^2}\sum_{i=1}^{N_j}\mathbb{E}\left[Y_i^2|\mathcal{F}_j\right]\right] \\
&= \mathbb{E}\left[\int_D \frac{1}{N_j}\mathbb{E}\left[\left(\nabla f\left(u_j,\cdot\right) - \mathbb{E}\left[\nabla f(u_j,\cdot)|\mathcal{F}_j\right]\right)^2|\mathcal{F}_j\right]\mathrm{d}x\right] \\
&\leq \mathbb{E}\left[\int_D \frac{1}{N_j}\mathbb{E}\left[\left(\nabla f(u_j,\cdot)\right)^2|\mathcal{F}_j\right]\mathrm{d}x\right] \\
&= \frac{1}{N_j}\mathbb{E}\left[\|\nabla f(u_j,\cdot)\|^2\right] \\
&\leq \frac{2}{N_j}\mathbb{E}\left[\|\nabla f(u_j,\cdot) - \nabla f(u_\star,\cdot)\|^2 + \|\nabla f(u_\star,\cdot)\|^2\right] \qquad \text{[Lipschitz condition (17)]} \\
&\leq \frac{2L^2}{N_j}\mathbb{E}\left[\|u_j - u_\star\|^2\right] + \frac{2}{N_j}\mathbb{E}\left[\|\nabla f(u_\star,\cdot)\|^2\right].
\end{aligned}$$

Finally, we have that

$$\begin{aligned}\mathbb{E}[\langle u_j - u_\star, T_2\rangle] &= \mathbb{E}[\mathbb{E}\left[\langle u_j - u_\star, T_2\rangle|\mathcal{F}_j\right]] \\ &= \mathbb{E}[\langle u_j - u_\star, \mathbb{E}\left[T_2|\mathcal{F}_j\right]\rangle] \\ &= \frac{1}{N_j}\sum_{i=1}^{N_j}\mathbb{E}[\langle u_j - u_\star, \mathbb{E}\left[Y_i|\mathcal{F}_j\right]\rangle] \\ &= 0,\end{aligned}$$

and, similarly, $\mathbb{E}[\langle T_1, T_2\rangle] = \mathbb{E}[\mathbb{E}\left[\langle T_1, T_2\rangle|\mathcal{F}_j\right]] = \mathbb{E}[\langle T_1, \mathbb{E}\left[T_2|\mathcal{F}_j\right]\rangle] = 0$, which concludes the proof. □

We now consider the FE version of (29):

$$u_{j+1}^h = u_j^h - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^h(u_j^h, \omega)]\,, \tag{31}$$

with $\overrightarrow{\omega_j} := (\omega_j^{(1)}, \cdots, \omega_j^{(N_j)})$, and focus on the particular setting $(\tau_j, N_j) = (\tau_0/j, \overline{N})$.

Theorem 4. *Suppose that the assumptions of Lemma 6 hold and let u_j^h denote the j-th iterate of (31) with $\tau_0 > 1/l$, $\tau_j = \frac{\tau_0}{j}$ and $N_j = \overline{N}$, $\forall j$. Then we have*

$$\mathbb{E}[\|u_j^h - u_\star\|^2] \le D_1 j^{-1} + D_2 h^{2r+2}, \tag{32}$$

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

Proof. It follows from Theorem 3 that the factor c_j in (30) for the particular case of $(\tau_j, N_j) = (\tau_0/j, \overline{N})$ becomes:

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

Next we use the recursive formula (30) and set, as in Section 3, $u_\star^h$ to be the exact optimal control for the FE problem defined in (11). We emphasize that (11) has no approximation in the probability space. Setting $a_j = \mathbb{E}[\|u_j^h - u_\star^h\|^2]$ and $\beta_j = \frac{2\tau_j^2}{\overline{N}}\mathbb{E}[\|\nabla f(u_\star^h, \cdot)\|^2]$, from (30) applied to the sequence of Finite Element solutions $\{u_j^h\}_{j>0}$ we find

$$\begin{aligned}a_{j+1} &\le c_j a_j + \beta_j \\ &\le c_j c_{j-1} a_{j-1} + c_j\beta_{j-1} + \beta_j \\ &\le \cdots \\ &\le \underbrace{\Big(\prod_{i=1}^{j} c_i\Big)}_{=\kappa_j} a_1 + \underbrace{\sum_{i=1}^{j}\beta_i \prod_{l=i+1}^{j} c_l}_{=\mathcal{B}_j}. \end{aligned} \tag{33}$$

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

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

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

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

and $\kappa_j \lesssim j^{-\tau_0 l}$. For the second term $\mathcal{B}_j$ in (33) we have:

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

For the term K_{ij} we can proceed as before

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

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 $D_4 > 0$ independent of h and j:

$$a_{j+1} \leq D_3 j^{-\tau_0 l} a_1 + D_4 j^{-1}. \tag{34}$$

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

$$a_{j+1} \leq D_1 j^{-1}, \tag{35}$$

with D_1 possibly depending in $\|u_0^h - u_\star^h\|$. Finally splitting the error as

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

and using (16) in Corollary 1 to bound the second term, the claim follows. □

Algorithm 1 contains a detailed pseudo-code description of the SG method (31) with a fixed FE mesh size h applied to the OCP (11).

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 $j = 1, \ldots, j_{max}$ **do**
  sample $\overline{N}$ realizations $a_j^{(i)} = a(\cdot, \omega_j^{(i)})$, $i = 1, \ldots, \overline{N}$, of the random field;
  solve $\overline{N}$ state problems $\rightarrow y(a_j^{(i)}, u)$, $i = 1, \ldots, \overline{N}$, using FE on mesh h;
  solve $\overline{N}$ adjoint problems $\rightarrow p(a_j^{(i)}, u)$, $i = 1, \ldots, \overline{N}$, using FE on mesh h;
  $\widehat{\nabla J} = \beta u + \frac{1}{\overline{N}} \sum_{j=1}^{\overline{N}} p(a_j^{(i)}, u)$
  $u = u - \frac{\tau_0}{j} \widehat{\nabla J}$
end

Algorithm 1: Stochastic Gradient algorithm with fixed mesh size.

We conclude this section by analyzing the complexity of the method described in Algorithm 1.

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

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

Here, we recall that the state and adjoint problems can be solved, using a triangulation with mesh size h, in computational time $C_h = O(h^{-n\gamma})$, $\gamma \in [1, 3]$, and r is the degree of the polynomial FE space.

Proof. To achieve a tolerance $O(tol^2)$ for the error $\mathbb{E}[\|u_j^h - u_\star\|^2]$, we can equidistribute the precision tol^2 over the two terms in (32). 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 2. *We have investigated also other choices for (τ_j, N_j) in the SG method* (31)*. However, among the cases considered, we have found that $(\tau_j, N_j) = (\frac{\tau_0}{j}, \overline{N})$ leads to the best complexity. For example we have studied the SG with step-size $\tau_j = \tau_0/j$, $\tau_0 l > 1$ and increasing the MC sample size as $N_j \sim j^{\tau_0 l - 1}$ across iterations. With this choice the estimate in* (32) *becomes*

$$a_{j+1} \leq D_4 j^{-\tau_0 l} \log(j)\,, \tag{36}$$

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

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

which is analogous to that of the full optimization algorithm discussed in Section 5. The proof of the bound (36) *is detailed in Appendix B for completeness.*

Remark 3. *In general convex stochastic optimization problems, the constant l may be challenging to estimate in practice, which makes it difficult to fulfill the condition $\tau_0 > 1/l$. To bypass this difficulty, one could consider the so-called* Averaged Stochastic Gradient method *[44] instead, in which the step size $\tau_j = \tau_0/j^\eta$, $\eta \in (0, 1)$,*

MC with linear solver	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

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. We remark, however, that in our setting the constant l is directly related to the regularization parameter, namely $l = 2\beta$, so the need for averaging the control is not so compelling.

Table 1 summarizes the complexity results for the SG method in both the fixed sample size and increasing sample size regimes, as well as the complexity of a linear solver (e.g. CG) applied to the fully discretized OCP. There, the total work (W) to achieve a given tolerance (tol) is presented. We see from the table that the SG version with fixed sample size $\overline{N}$ (second column) improves the complexity only by a logarithmic factor compared to the MC method in conjunction with a linear solver for the optimization problem (first column). The advantage we see in this SG version compared to the MC method with linear solver is that we do not have to fix the sample size N in advance and can, instead, simply monitor the convergence of the SG iterations until a prescribed tolerance is reached. However, in Algorithm 1, we do have to choose the FE mesh size in advance. 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 discus a variant of the stochastic gradient (SG) method that also refines the FE mesh size across iterations of the steepest decent optimization routine. That is, the new mesh size h_j is now depending on the current 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 a non-decreasing function. The complete SG procedure with decreasing FE mesh size then reads:

$$u_{j+1}^{h_{j+1}} = u_j^{h_j} - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \cdot)], \tag{37}$$

with $\overrightarrow{\omega_j} := (\omega_j^{(1)}, \cdots, \omega_j^{(N_j)})$. Notice that if non-nested meshes are used across the iterations, a projection operator has to be added in (37) to transfer information from one mesh to another. We first derive a recurrence formula for the error in the spirit of (30).

Theorem 5. *Let $u_{j+1}^{h_{j+1}}$ be the approximated control obtained from the SG with variable mesh size* (37) *and $u_\star$ the exact control for the continuous optimal problem* (10)*. Then, under the assumptions of Lemma 6, we have:*

$$\begin{aligned}\mathbb{E}[\|u_{j+1}^{h_{j+1}} - u_\star\|^2] \\ \leq c_j \mathbb{E}[\|u_j^{h_j} - u_\star\|^2] + \frac{4\tau_j^2}{N_j}\mathbb{E}[\|\nabla f(u_\star, \cdot)\|^2] + 4\tau_j\big(\tau_j(1 + \frac{2}{N_j}) + \frac{1}{l}\big)C h_j^{2r+2},\end{aligned} \tag{38}$$

where $c_j = 1 - \frac{\tau_j l}{2} + \tau_j^2 L^2\big(2 + \frac{2}{N_j}\big)$, l *and* L *are the convexity constant and the Lipschitz constant of* f*, resp., and* $C > 0$ *is a constant that depends on the* H^{r+1}*-seminorm of* $y(u_\star)$ *and* $p(u_\star)$*.*

Proof. Subtracting the optimal continuous control $u_\star$ from both sides of the recurrence formula (37), we get

$$\begin{aligned}
u_{j+1}^{h_{j+1}} - u_\star =& u_j^{h_j} - u_\star - \tau_j E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \cdot)] \pm \tau_j \mathbb{E}[\nabla f^{h_j}(u_\star, \cdot)] \\
& \pm \tau_j \mathbb{E}\left[\nabla f^{h_j}(u_j^{h_j}, \cdot) | \mathcal{F}_j\right] + \tau_j \mathbb{E}[\nabla f(u_\star, \cdot)] \\
=& u_j^{h_j} - u_\star + \tau_j \left(\mathbb{E}[\nabla f^{h_j}(u_\star, \cdot)] - \mathbb{E}\left[\nabla f^{h_j}(u_j^{h_j}, \cdot) | \mathcal{F}_j\right]\right) \\
& + \tau_j \left(\mathbb{E}\left[\nabla f^{h_j}(u_j^{h_j}, \cdot) | \mathcal{F}_j\right] - E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \cdot)]\right) + \tau_j \left(\mathbb{E}[\nabla f(u_\star, \cdot) - \nabla f^{h_j}(u_\star, \cdot)]\right).
\end{aligned}$$

Then setting, similar as in the proof of Theorem 3,

$$\begin{aligned}
T_1 &:= \mathbb{E}\left[\nabla f^{h_j}(u_j^{h_j}, \cdot) | \mathcal{F}_j\right] - \mathbb{E}[\nabla f^{h_j}(u_\star, \cdot)], \\
T_2 &:= E_{MC}^{\overrightarrow{\omega_j}}[\nabla f^{h_j}(u_j^{h_j}, \cdot)] - \mathbb{E}\left[\nabla f^{h_j}(u_j^{h_j}, \cdot) | \mathcal{F}_j\right], \\
T_3 &:= \mathbb{E}[\nabla f(u_\star, \cdot) - \nabla f^{h_j}(u_\star, \cdot)],
\end{aligned}$$

we can rewrite the last equality as

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

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

$$\begin{aligned}
\mathbb{E}[\|u_{j+1}^{h_{j+1}} - u_\star\|^2] =& \mathbb{E}[\|u_j^{h_j} - u_\star\|^2] + \tau_j^2 \mathbb{E}[\|T_1\|^2] + \tau_j^2 \mathbb{E}[\|T_2\|^2] + \tau_j^2 \mathbb{E}[\|T_3\|^2] \\
& - 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u_\star, T_1\rangle] - 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u_\star, T_2\rangle] + 2\tau_j \mathbb{E}[\langle u_j^{h_j} - u_\star, T_3\rangle] \\
& + 2\tau_j^2 \mathbb{E}[\langle T_1, T_2\rangle] - 2\tau_j^2 \mathbb{E}[\langle T_2, T_3\rangle] - 2\tau_j^2 \mathbb{E}[\langle T_1, T_3\rangle] \,.
\end{aligned} \tag{39}$$

Next, we bound each of these ten terms on the right-hand side. First, the term $\tau_j^2 \mathbb{E}[\|T_1\|^2]$ can be bounded as in the proof of Theorem 3 leading to:

$$\tau_j^2 \mathbb{E}[\|T_1\|^2] \le \tau_j^2 L_{h_j}^2 \mathbb{E}[\|u_j^{h_j} - u_\star\|^2],$$

with L_{h_j} being the Lipschitz constant for the function f^{h_j}, which is bounded by L (see Lemma 7). For the term $\tau_j^2 \mathbb{E}[\|T_3\|^2]$, we find,

$$\begin{aligned}
\tau_j^2 \mathbb{E}[\|T_3\|^2] &= \tau_j^2 \|\mathbb{E}[\nabla f(u_\star, \cdot) - \nabla f^{h_j}(u_\star, \cdot)]\|^2 \\
&= \tau_j^2 \|\mathbb{E}[p(u_\star) - p^{h_j}(u_\star)]\|^2 \\
&\le \tau_j^2 \mathbb{E}[\|p(u_\star) - p^{h_j}(u_\star)\|^2] \\
&\le 2\tau_j^2 \mathbb{E}[\|p(u_\star) - \widetilde{p}^{h_j}(u_\star)\|^2] + 2\tau_j^2 \mathbb{E}[\|\widetilde{p}^{h_j}(u_\star) - p^{h_j}(u_\star)\|^2] \\
&\le 2C\tau_j^2 \mathbb{E}[|p(u_\star)|_{H^{r+1}}^2] h^{2r+2} + 2C\tau_j^2 \mathbb{E}[|y(u_\star)|_{H^{r+1}}^2] h^{2r+2} \quad \text{[using Lemma 6]} \\
&\le 2\tau_j^2 C(y(u_\star), p(u_\star)) h^{2r+2}.
\end{aligned}$$

Next, for $\tau_j^2 \mathbb{E}[\|T_2\|^2]$ we use the same steps as in Theorem 3 to find

$$\tau_j^2 \mathbb{E}[\|T_2\|^2] \le \frac{2\tau_j^2 L_{h_j}^2}{N_j} \mathbb{E}\Big[\|u_j^{h_j} - u_\star\|^2\Big] + \frac{2\tau_j^2}{N_j} \mathbb{E}\Big[\|\nabla f^{h_j}(u_\star, \cdot)\|^2\Big].$$

The second term of the right hand side can be further bounded uniformly w.r.t. h_j as

$$\begin{aligned}\mathbb{E}[\|\nabla f^{h_j}(u_\star,\cdot)\|^2] &\le 2\mathbb{E}[\|\nabla f^{h_j}(u_\star,\cdot)-\nabla f(u_\star,\cdot)\|^2]+2\mathbb{E}[\|\nabla f(u_\star,\cdot)\|^2]\\ &\le 4C(y(u_\star),p(u_\star))h_j^{2r+2}+2\mathbb{E}[\|\nabla f(u_\star,\cdot)\|^2],\end{aligned}$$

where we have used the same steps as for T_3 to bound $\mathbb{E}[\|\nabla f^{h_j}(u_\star,\cdot)-\nabla f(u_\star,\cdot)\|^2]$. Finally, for the cross terms we have

$$\begin{aligned}2\tau_j\mathbb{E}[\langle u_j^{h_j}-u_\star,T_1\rangle] &= 2\tau_j\mathbb{E}[\langle u_j^{h_j}-u_\star,\mathbb{E}\left[\nabla f^{h_j}(u_j^{h_j},\cdot)-\nabla f^{h_j}(u_\star,\cdot)|\mathcal{F}_j\right]\rangle]\\ &= 2\tau_j\mathbb{E}[\mathbb{E}\left[\langle u_j^{h_j}-u_\star,\nabla f^{h_j}(u_j^{h_j},\cdot)-\nabla f^{h_j}(u_\star,\cdot)\rangle|\mathcal{F}_j\right]]\\ &\ge \tau_j l\mathbb{E}[\|u_j^{h_j}-u_\star\|^2],\quad\text{[using Strong convexity]}\end{aligned}$$

and as in Theorem 9,

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

Moreover

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

and finally

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

Putting everything together, we finally obtain (38), 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 (38). This leads to the following.

Theorem 6. *Suppose that the assumptions of Lemma 6 hold and let $u_j^{h_j}$ denote the j-th iterate of the SG method with variable mesh size* (37)*. For the particular choice $(\tau_j,N_j,h_j)=(\tau_0/j,\overline{N},h_0 2^{-\ell(j)})$, with $\ell(j)=\left\lceil\frac{\ln_2(j)}{2r+2}\right\rceil$, and assuming $\tau_0>1/l$, we have:*

$$\mathbb{E}[\|u_j^{h_j}-u_\star\|^2]\le F_1 j^{-1} \tag{40}$$

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 $\frac{4\tau_j^2}{N_j}\mathbb{E}[\|\nabla f_{h_j}(u_\star,\cdot)\|^2]$ and $4\tau_j\big(\tau_j(1+\frac{2}{N_j})+\frac{1}{l}\big)Ch_j^{2r+2}$ in the inequality (38) have the same order $O(j^{-2})$. Then, we apply the same reasoning as in Theorem 4 to conclude the proof. □

Algorithm 2 details the SG Robbins–Monro method with variable mesh size (37) applied to the OCP (10).

Data:
Given a desired tolerance tol, choose $\tau_0 > \frac{1}{l}$, h_0 and $j_{max} \simeq tol^{-2}$
initialization:
$u = 0$
for $j = 1, \dots, j_{max}$ **do**
  update mesh size to $h = h_0 2^{-\lceil \frac{\ln_2 j}{2r+2} \rceil}$
  sample $\overline{N}$ realizations $a_j^{(i)} = a(\cdot, \omega_j^{(i)})$, $i = 1, \dots, \overline{N}$, of the random field
  solve $\overline{N}$ state problems $\rightarrow y(a_j^{(i)}, u)$, $i = 1, \dots, \overline{N}$, on mesh h
  solve $\overline{N}$ adjoint problems $\rightarrow p(a_j^{(i)}, u)$, $i = 1, \dots, \overline{N}$, on mesh h
  $\widehat{\nabla J} = \beta u + \frac{1}{\overline{N}} \sum_{i=1}^{\overline{N}} p(a_j^{(i)}, u)$
  $u = u - \frac{\tau_0}{j} \widehat{\nabla J}$
end

Algorithm 2: Stochastic Gradient algorithm with variable mesh size.

Concerning the complexity of the Algorithm 2, one can derive the following complexity result.

Corollary 3. *With the choice* $(\tau_j, N_j, , h_j) = (\tau_0/j, \overline{N}, h_0 2^{-\ell(j)})$ *as in Theorem 6, in order to achieve a given tolerance* $O(tol)$ *in a root mean squared sense, i.e. to guarantee that* $\mathbb{E}[\|u_j^{h_j} - u_\star\|^2] \lesssim tol^2$*, the total required computational work* W *is bounded by:*

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

Proof. To guarantee that the mean squared error tolerance is met, in view of Theorem 6 it suffices to ensure that $tol^2 \lesssim j_{max}^{-1}$, which holds for $j_{max} \simeq tol^{-2}$. Then the total work required is bounded by

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

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

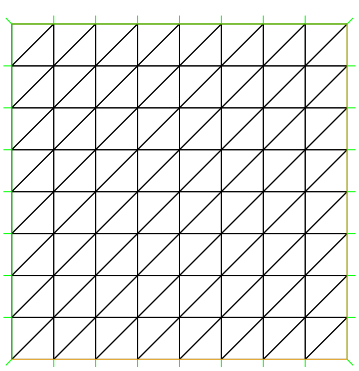
$$\begin{aligned} W &\leq 2\overline{N} \sum_{p=1}^{j_{max}} 2^{n\gamma \left(\frac{\ln_2 p}{2r+2} + 1 \right)} \leq 2^{n\gamma+1} \overline{N} \sum_{p=1}^{j_{max}} p^{\frac{n\gamma}{2r+2}} \\ &\leq 2^{n\gamma+1} \overline{N} \frac{2r+2}{2r+2+n\gamma} \left(j_{max} + 1 \right)^{\frac{n\gamma}{2r+2}+1}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

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

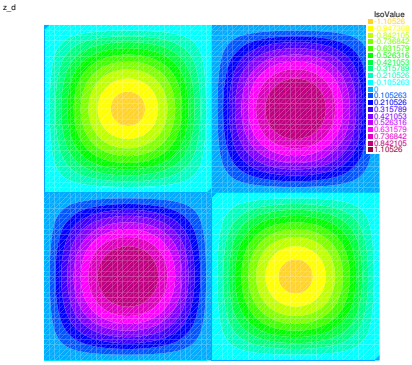
$$W \lesssim tol^{-2-\frac{n\gamma}{r+1}}.$$

□

Notice that the asymptotic complexity remains the same as for the Stochastic Gradient algorithm with fixed mesh size (31); cf. Corollary 2. However, as the SG method with variable mesh size uses computations on coarser meshes for the first few iterations, we nonetheless expect a practical improvement due to reducing the proportionality constant. We will assess such improvement by numerical experimentation in the next Section.



(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.

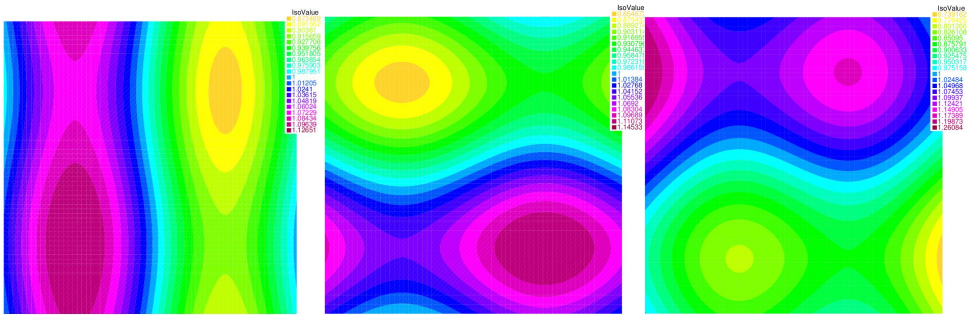
8. Numerical results

In this section we verify the assertions of Theorems 2, 4, and 6, as well as the computational complexity derived in the corresponding Corollaries 1, 2 and 3. Specifically, we illustrate the order of convergence for the three versions of the iterative optimization method presented in Sections 5, 6, and 7 respectively. For this purpose, we consider problem (2) in the domain $D = (0,1)^2$ with $g = 1$ and the random diffusion coefficient

$$a(x_1, x_2, \boldsymbol{\xi}) = 1 + \exp\left(var\left(\xi_1 \cos(1.1\pi x_1) + \xi_2 \cos(1.2\pi x_1) + \xi_3 \sin(1.3\pi x_2) + \xi_4 \sin(1.4\pi x_2)\right)\right), \tag{41}$$

with $(x_1, x_2) \in D$, $var = \exp(-1.125)$ and $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_4)$ with $\xi_i \overset{i.i.d.}{\sim} \mathcal{U}([-1,1])$. Figure 2 shows three typical realizations of the random field. The target function z_d has been chosen as $z_d(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$ (see Fig. 1 b) and we have taken $\beta = 10^{-4}$ 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 finite elements (i.e. $r = 1$). All computations have been performed using the FE library Freefem++[26]. This relatively simple setting, with only 4 uniform random variables, has been chosen to be able to compute an accurate reference solution by a stochastic collocation method on a fine FE mesh.

Reference solution. To compute a reference solution of problem (2), we use a full tensorized Gauss-Legendre (GL) quadrature formula with 9 points in each direction (i.e. with a total number of knots $N = 9^4$) and a fine triangulation with $h = 2^{-9}$; see, e.g., references [8, 43] and Appendix A for estimates of the quadrature error. As this approximated problem with fixed Gauss nodes is now deterministic, we have used a stopping condition based on the norm of the gradient, and have chosen the conjugate gradient (CG) algorithm applied on the linear system (24) as the iterative optimization scheme. In Figure 3 we show the optimal control obtained after $j = 16$ iterations when the stopping criterion $\|E^{GL}_{(9,9,9,9)}[\nabla J(u^h_j)]\| \leq 10^{-8}$ was met, where u^h_j is the j-th CG iterate and $\widehat{E} = E^{GL}_{(9,9,9,9)}$ denote the tensor Gaussian quadrature used in (19) to approximate the true expectation. The L^2-norm of the final control using this Gaussian quadrature is $\|\widehat{u}^{h=2^{-8}}_{j=16}\| = 16.4128$.



(A) $\xi_1 = 0.0327973$
$\xi_2 = 0.10508$
$\xi_3 = 0.141335$
$\xi_4 = 0.905369$

(B) $\xi_1 = 0.370554$
$\xi_2 = 0.0682218$
$\xi_3 = 0.667794$
$\xi_4 = -0.421315$

(C) $\xi_1 = -0.943052$
$\xi_2 = 0.968895$
$\xi_3 = -0.656957$
$\xi_4 = -0.997339$

FIGURE 2. Three realizations of the diffusion random field (41).

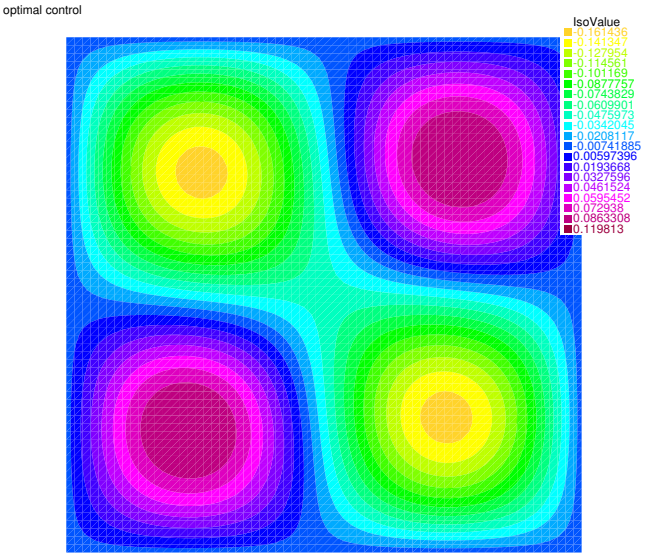


FIGURE 3. Optimal control reference solution computed with $h = 2^{-8}$ on tensorized Gauss-Legendre quadrature formula with $N = 9^4$ nodes.

8.1. **Conjugate gradient on fully discretized OCP.** We investigate here the convergence of the method described in Section 5, with the particular choice of the CG method applied to the linear system (24) as iterative solver. We recall the error bound (27) in the case of piece-wise linear FE (i.e. $r = 1$):

$$\mathbb{E}[\|\widehat{u}_j^h - u\|^2] \leq C_1 e^{-\rho j} + \frac{C_2}{N_{MC}} + C_3 h^4 \,. \tag{42}$$

For each tolerance tol, using formula (42), we compute the optimal mesh size $h = h(tol)$, the optimal sample size in the MC approximation $N_{MC} = N_{MC}(tol)$, and, finally, the minimum number of iterations required in the iterative optimization method, $j = j(tol)$. To optimally balance the error contributions in (42), we need further to estimate the constants C_1, C_2, C_3, ρ. This is detailed hereafter.

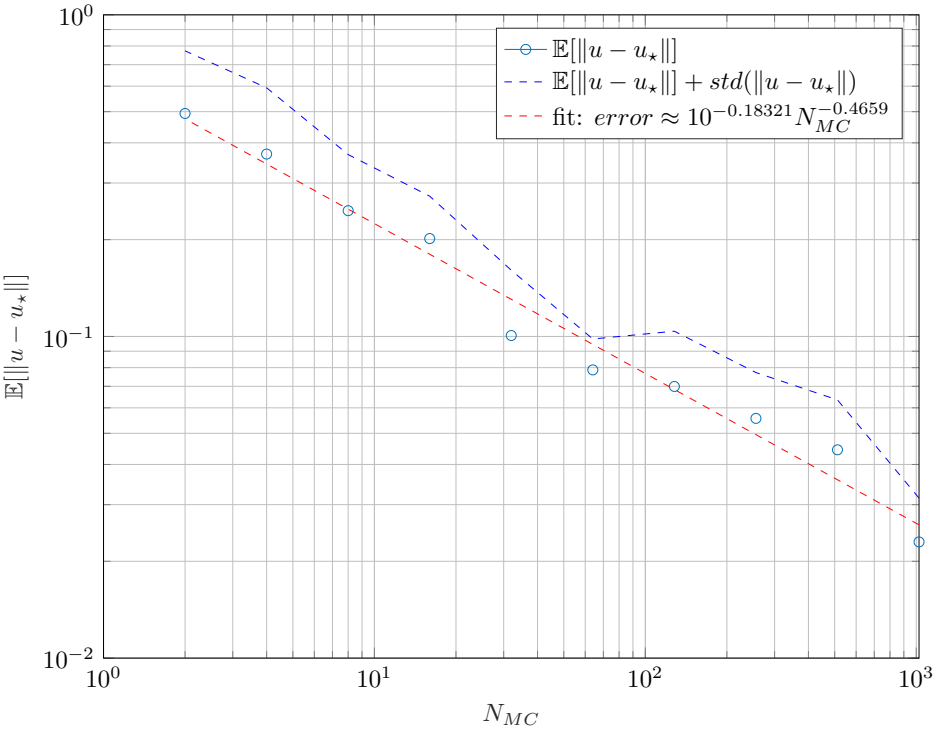


FIGURE 4. L^2-error on the optimal control of the fully discretized OCP with mesh size $h = 2^{-9}$ and increasing MC sample size N_{MC}. Mean and std of the error estimated by sample averages over 10 independent realizations.

- In order to estimate the constant C_2, we used the same finest mesh as the one used to compute our reference solution, namely $h = 2^{-9}$, and ran the CG method up to 20 iterations on the linear system (24) discretized by Monte Carlo with a sample of increasing size $N_{MC} = 2^0, 2^1, \cdots, 2^{10}$. For every sample size N_{MC} we repeated the simulation 10 times (with 10 independent MC samples) and averaged the final $L^2(D)$ error on the control. We numerically found $C_2 \approx 0.430527$. Figure 4 presents these results, where the mean and standard deviation (std) of the L^2 error have been approximated by sample averages using the 10 independent realizations.
- To estimate the constant C_3, we have discretized the OCP by a Gaussian quadrature with $N = 2^4$ knots (2 Gauss-Legendre points per random variable) and a sequence of decreasing mesh sizes $h = 2^{-1}, \ldots, 2^{-8}$. The optimal control has been computed by sufficiently many CG iterations and the error estimated with respect to a reference solution with $h = 2^{-9}$ and the same Gaussian quadrature with $N = 2^4$ knots. We found $C_3 \approx 436.516$. Figure 5 shows the convergence of the error on the control (in the L^2-norm), versus the discretization parameter h. We observe a convergence rate of h^2, which is consistent with the theoretical result in (42).
- To estimate the constants C_1 and ρ, we have discretized the OCP by a Gaussian quadrature with $N = 5^4$ tensorized Gauss-Legendre knots and $h = 2^{-9}$. We have run the CG algorithm and recorded the L^2 error on the control, computed with respect to a converged solution, over the iterations. We found $\rho = 4.941$ and $C_1 = 17.9140$.

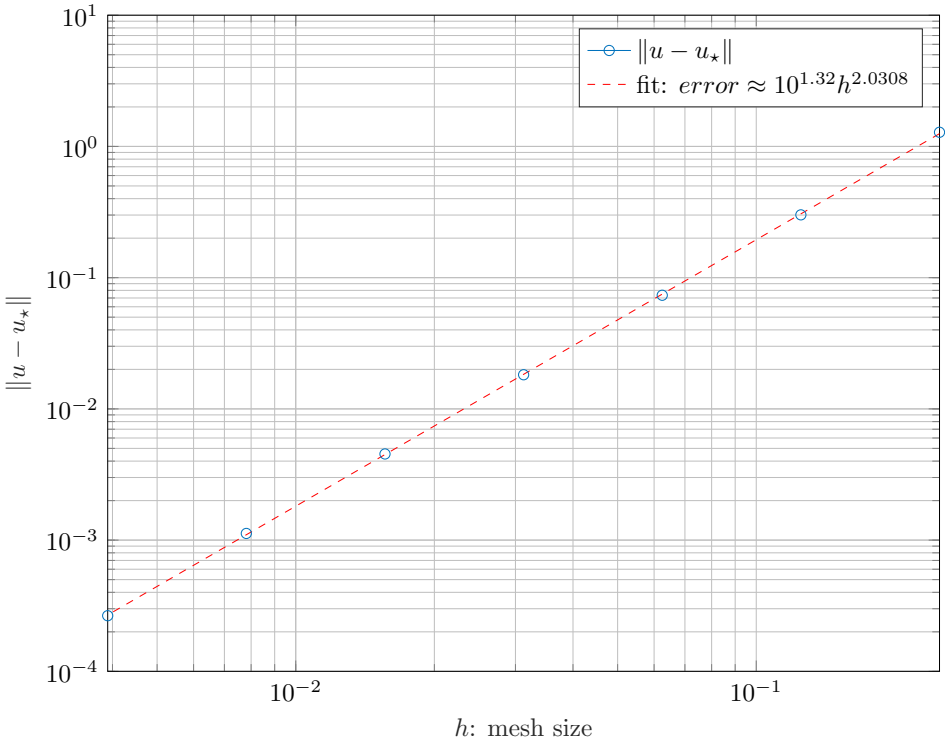


FIGURE 5. L^2-error on the optimal control of the fully discrete OCP with $N = 2^4$ Gauss-Legendre quadrature knots and decreasing mesh sizes.

Figure 6 shows the computational complexity of the considered method, i.e. the fully discretized OPC solved by the CG algorithm, with optimally chosen parameters N_{MC}, h, and j. Here we plot the mean L^2 error on the optimal control versus the computational cost model $W = 2N_{MC}j(\frac{1}{h} - 1)^2$ (which assumes an optimal linear algebra solver with $\gamma = 1$ for the FE discretized PDE). The mean error and its standard deviation have been estimated by repeating the whole procedure 20 times. 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.** We implement here the Stochastic Gradient method described in Section 6 using $\overline{N} = 1$ samples at each iteration (recall that the complexity does not depend on $\overline{N}$) and learning rate $\tau_j = \frac{\tau_0}{j+10}$, with $\tau_0 = \frac{2}{\beta}$. We have first assessed the convergence of the SG iterations on the FE discretized OCP (11), using a mesh size $h = 2^{-4}$. The reference solution was computed using the same FE mesh size, a Gaussian quadrature with $N = 5^4$ knots to approximate the expectation and CG iterations up to convergence. Figure 7 clearly shows the $\frac{1}{\sqrt{j}}$ convergence rate of the SG algorithm w.r.t. the iteration counter j, as predicted by Theorem 4.

We have then studied the complexity of the SG Algorithm (where the error is computed with respect to the solution of the continuous OCP (10)). For convenience, we recall the error bound (32):

$$\mathbb{E}[\|u_j^h - u_\star\|^2] \leq D_1 j^{-1} + D_2 h^{2r+2}.$$

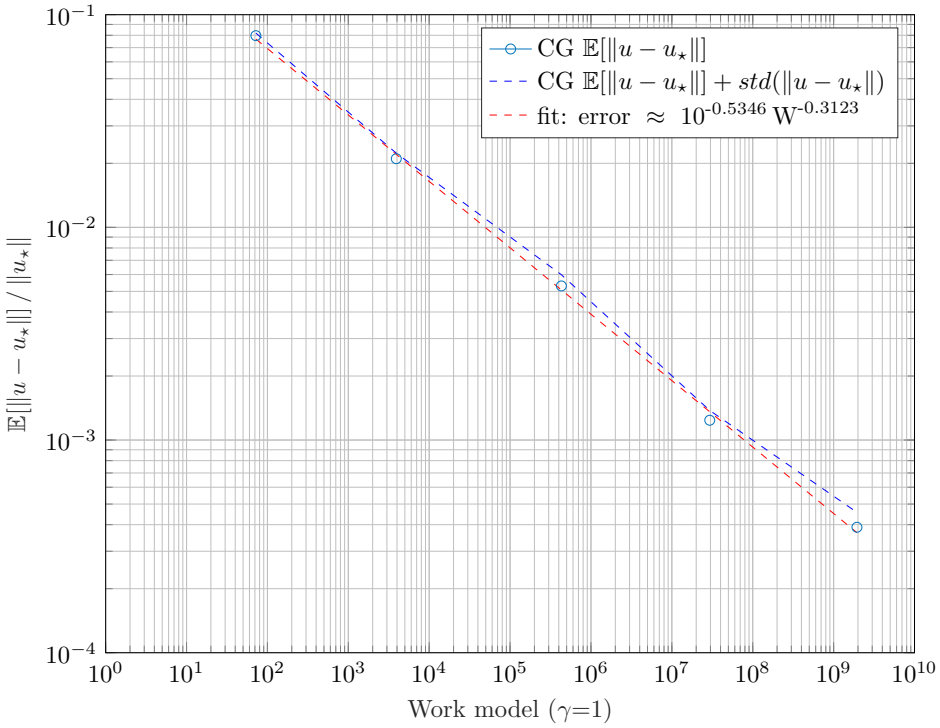


FIGURE 6. Mean relative L^2-error on the optimal control versus the computational work model $W = 2N_{MC}j(\frac{1}{h}-1)^2$ for the fully discretized OPC solved by the CG algorithm (with optimally chosen parameters N_{MC}, h, j). Mean and standard deviation of the error estimated over 20 repetitions.

To optimally balance the two error contributions, we have estimated the constants D_1 and D_2 as described next.

- The constant D_1 can be inferred from the results in Figure 7. A least squares fit of the mean L^2 error versus the iteration counter j gives $D_1 \approx 2.143$.
- The constant D_2 is the same as the constant C_3 in (42), hence we kept the same estimate $D_2 \approx 436.5$.

With these constants estimated, for a given required tolerance tol we can estimate the correct number of SG iterations $j(tol)$ and mesh size $h(tol)$ to fulfill the accuracy requirement. Figure 8 shows the estimated mean L^2 error, using SG Algorithm 1, as a function of the computational cost model $W = 2j(\frac{1}{h}-1)^2$. The slope is the one predicted in Corollary 2 (with $r = 1$ and $\gamma = 1$), namely $W \lesssim tol^{-3}$.

8.3. **Stochastic Gradient with variable mesh size.** We present here the results for the Stochastic Gradient method described in Section 7 with $\bar{N} = 1$ and learning rate $\tau_j = \frac{\tau_0}{j+15}$, with $\tau_0 = \frac{2}{\beta}$. The mesh refinement strategy over the iterations is the one described in Algorithm 2 with $h_0 = 2^{-4}$.

The rational behind this choice of h_0 is the following: from (38) we have that

$$\mathbb{E}[\|u_{j+1}^{h_{j+1}} - u_\star\|^2] \leq c_j \mathbb{E}[\|u_j^{h_j} - u_\star\|^2] + MC + FE$$

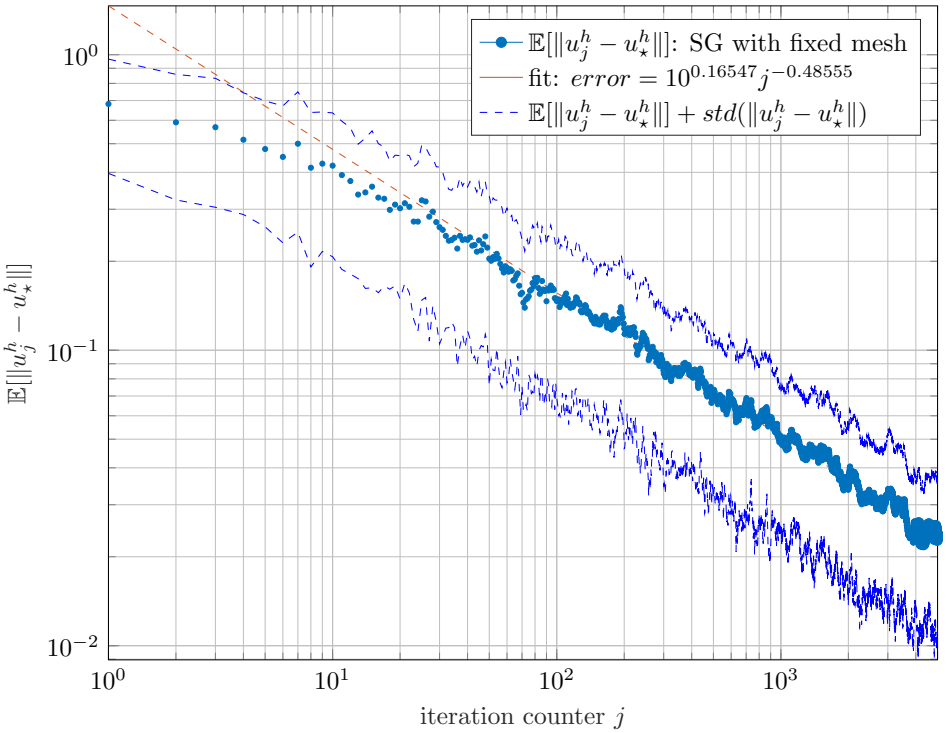


FIGURE 7. SG Algorithm applied to the FE discretized OPC (11) with $h = 2^{-4}$. Mean L^2 error as a function of iteration counter, estimated by sample average over 100 independent realizations.

where the Monte Carlo error behaves asymptotically for $j \to \infty$ as $MC = \frac{4\tau_j^2}{N}\mathbb{E}[\|\nabla f(u_\star, \cdot)\|^2] \sim \frac{4\tau_0^2}{j^2}C_2$ with $C_2 \approx 0.43$ estimated in Section 8.1, and the Finite Element error behaves as $FE = 4\tau_j\big(\tau_j(1 + \frac{2}{N}) + \frac{1}{l}\big)C_3 h_j^{2r+2} \sim \frac{4\tau_0(3\tau_0 + \frac{1}{l})}{j^2}C_3 h_0^4$, with $C_3 \approx 436.5$ estimated in Section 8.1. Equilibrating the two error contributions MC and FE, leads to an h_0 of the order of the chosen one .

We have run the algorithm for $j_{max} = 10000$ iterations, and repeated the simulation 100 times to estimate the mean error and its standard deviation.

Figure 9 shows the mean L^2 error computed with respect to the same reference solution described at the beginning of this section, versus the iteration counter. Figure 10 shows, instead, the mean L^2 error at iteration j versus the computational cost model $W_j = \sum_{k=1}^{j} 2(\frac{1}{h_k} - 1)^2$, as well as one particular realization of the algorithm. In both plots, the observed convergence rate of the mean error is consistent with the results in Theorem 6 and Corollary 3, resp. In Figure 10 we have also added, for comparison purposes, the complexity results of the other two methods, namely CG on the fully discretized OCP and SG on the semi-discretized OCP, with optimal choices of discretization parameters. It is clear from this plot that all methods perform very similarly for the problem at hand.

9. Conclusions

In this work, we have analyzed and compared the complexity of three gradient based methods for the numerical solution of a risk-averse optimal control problem involving an elliptic PDE with random coefficients, where a Finite Element discretization is used to approximate the underlying PDEs and a Monte Carlo

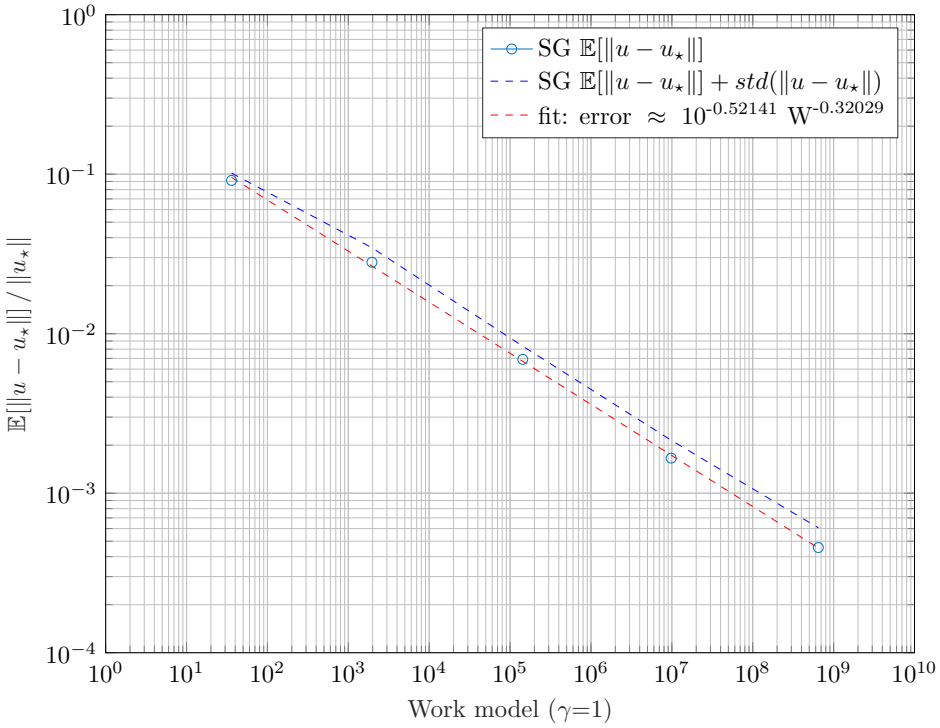


FIGURE 8. Mean relative L^2 error on the optimal control versus the computational work model $W = 2j(\frac{1}{h}-1)^2$ for the SG Algorithm 1 (with optimally chosen parameters j and h). Mean and standard deviation of the error estimated over 20 repetitions.

sampling is used to approximate the expectation in the risk measure. The first version considered is when the OCP is discretized upfront, using a fixed finite element mesh and a fixed Monte Carlo sample, and then solved by an efficient iterative method such as a Conjugate Gradient. The second version is a Stochastic Gradient method in which the finite element discretization is still kept fixed over the iterations, but 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 gradient methods improve the computational complexity by log factors, compared to applying the CG (or equivalent) linear solver to the fully discretized OCP. Our complexity analysis is based on a-priori error estimates and a-priori choices for the FE mesh size, the Monte Carlo sample size, and the gradient iterations to obtain a prescribed tolerance.

In addition to the improved complexity, another benefit of the stochastic gradient methods is that they are more amenable to adaptive versions, in which, e.g., 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.

Another interesting direction for future work is the extension of stochastic gradient methods to more general risk measures. For example, 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

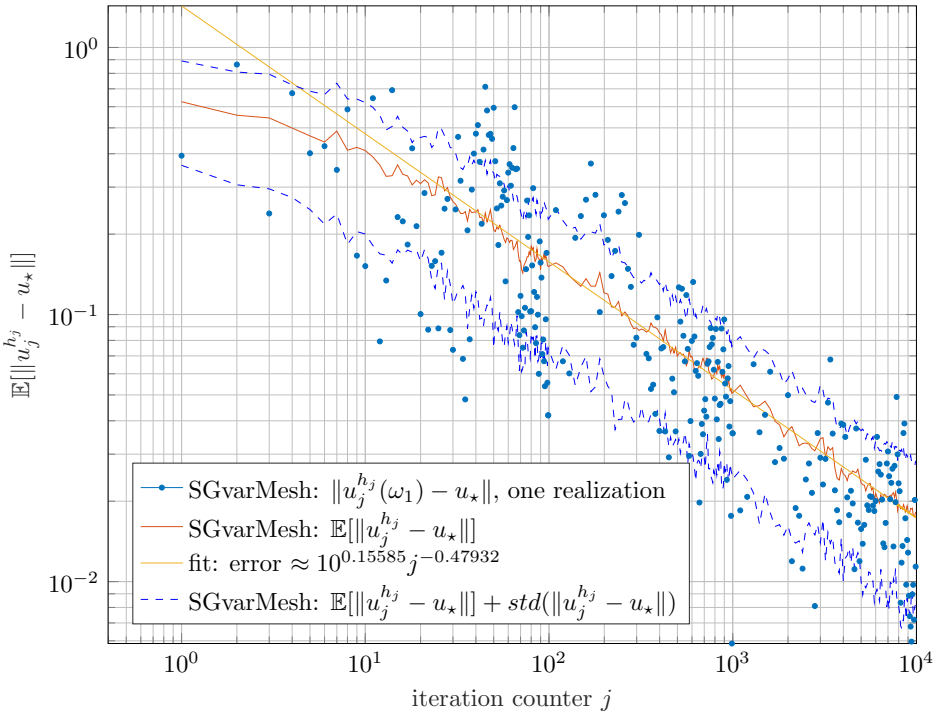


FIGURE 9. SG Algorithm with variable mesh size, and initial step-size $\tau_0 = 2/\beta$. Mean L^2 error $\mathbb{E}[\|u - u_\star\|]$, averaged over 100 experiments, as a function of the iteration counter j.

problems. One way of introducing more general risk mesures in our context could be to consider the objective functional

$$J(u) = \frac{1}{2}\int_D \sigma\Big(y_\omega(u)(x) - z_d(x)\Big)dx + \frac{\beta}{2}\|u\|^2.$$

In this work we have focuced on the "mean-squared" risk measure $\sigma(\cdot) = \mathbb{E}[(\cdot)^2]$ but this could be replaced by any other coherent and more risk averse measure.

ACKNOWLEDGMENTS

The authors have received support from the Center for ADvanced MOdeling Science (CADMOS). The third author also acknowledges funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 800898, project ExaQUte — EXAscale Quantification of Uncertainties for Technology and Science Simulation, as well as from the Swiss National Science Foundation under the Project n. 172678 "Uncertainty Quantifi-cation techniques for PDE constrained optimization and random evolution equations".

APPENDIX A. REFERENCE SOLUTION BY STOCHASTIC COLLOCATION

In this appendix, we briefly describe the computation of the reference solution used in the numerical result of Section 8, which is based on a stochastic collocation method on a tensor grid of Gauss Legendre points. Moreover, we provide an error estimate for such an approximation based on stochastic collocation. While the numerical example in Section 8 only depends on 4 random variables, here we show that the stochastic collocation approximation is exponentially convergent and

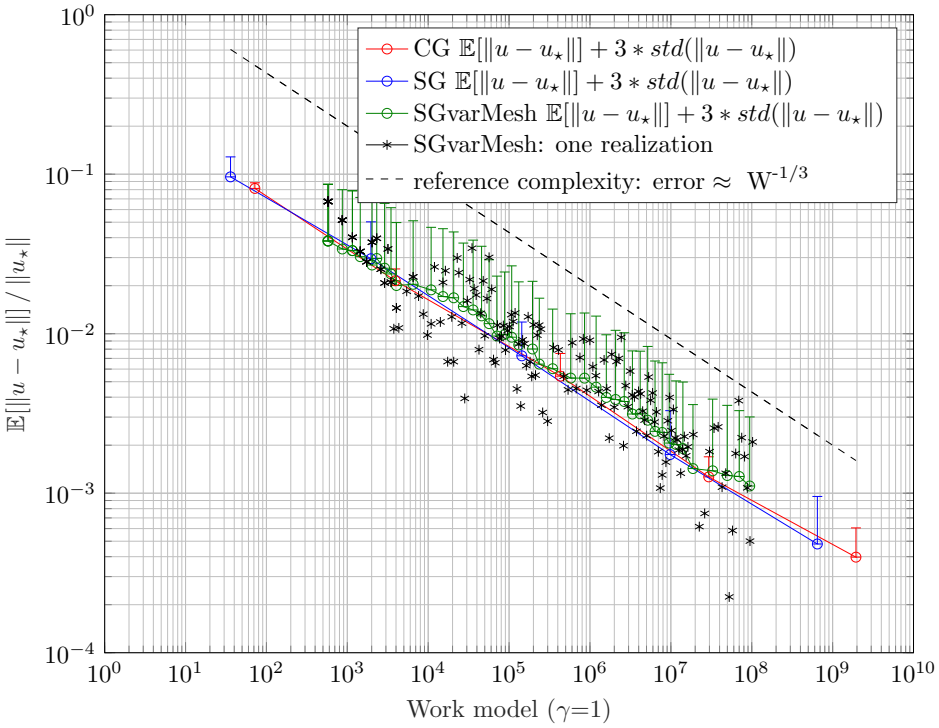


FIGURE 10. Comparison between the three different solution methods: CG with fixed mesh and MC sample; SG with fixed mesh, SG with variable mesh. The estimated mean relative L^2 error $\mathbb{E}[\|u-u_\star\|]$ is plotted as a function of the theoretical computational work W. Upper error bar corresponds to 3 standard deviations. For the latter method, also one particular realization is shown.

that a highly accurate solution can be obtained with a moderate number of collocation points; recall that 9^4 points were used in the numerical experiments. We suppose that quadrature to approximate the expectation is not random, but uses deterministic points ξ_i, for $i = 1, \dots, N$. The estimated optimal control $\widehat{u}$ is then deterministic as well. The following theorem then provides an error bound.

Theorem 7. *Denoting by $u_\star$ the optimal control solution of the exact problem* (10) *and by $\widehat{u}$ the solution of the semi-discrete collocation problem* (19) *using a deterministic quadrature formula $\widehat{E}$ as approximation for $\mathbb{E}$, we have*

$$\frac{\beta}{2}\|\widehat{u}-u_\star\|^2 + \mathbb{E}[\|y(u_\star)-y(\widehat{u})\|^2] \le \frac{1}{2\beta}\|\mathbb{E}[p(\widehat{u})]-\widehat{E}[p(\widehat{u})]\|^2. \tag{43}$$

Proof. The expressions of the gradient of J and $\widehat{J}$ are given by $\nabla J(u_\star) = \beta u_\star + \mathbb{E}[p(u_\star)]$, $\nabla \widehat{J}(\widehat{u}) = \beta\widehat{u} + \widehat{E}[p(\widehat{u})]$, respectively, and the corresponding optimality conditions read

$$\langle \nabla J(u_\star), v\rangle = 0, \quad \langle \nabla \widehat{J}(\widehat{u}), v\rangle = 0, \quad \forall v \in U. \tag{44}$$

Then choosing $v = \widehat{u} - u_\star$ in (44), and combining both we find

$$\langle \beta(u_\star - \widehat{u}) + \mathbb{E}[p(u_\star)] - \widehat{E}[p(\widehat{u})], \widehat{u} - u_\star\rangle = 0,$$

that is,

$$\beta\|u_\star - \widehat{u}\|^2 = \langle \mathbb{E}[p(u_\star)] - \mathbb{E}[p(\widehat{u})] + \mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})], \widehat{u} - u_\star\rangle. \tag{45}$$

In order to bound the first part of the error in (45), $\langle \mathbb{E}[p(u_\star)] - \mathbb{E}[p(\widehat{u})], \widehat{u} - u_\star\rangle$, we can write for any $\omega \in \Omega$

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

Then, taking expectation, we find

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

For the second contribution, $\langle \mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})], \widehat{u} - u_\star\rangle$, Cauchy-Schwarz and Young's inequalities yield

$$\langle \mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})], \widehat{u} - u_\star\rangle \le \frac{1}{2\beta}\|\mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})]\|^2 + \frac{\beta}{2}\|\widehat{u} - u_\star\|^2,$$

from which the claim follows. □

The quantification of the quadrature error $\mathbb{E}[p(\widehat{u})] - \widehat{E}[p(\widehat{u})]$, i.e., the right hand side in (43), heavily depends on the smoothness of the adjoint function in the stochastic variables. The numerical example considered in Section 8 has a diffusion coefficient that is a particular version of the more general class of diffusion coefficients 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, $\|b_i\|_{L^\infty(D)} = 1$, $\sum_{i=1}^{M}\sqrt{\lambda_i} < \operatorname{ess\,inf}_{x\in D} a_0(x)$, and $\xi_i \sim \mathcal{U}([-1,1])$ i.i.d. 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. Moreover, as $\xi_i \sim \mathcal{U}([-1,1])$ here, 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 $(k_1, \ldots, k_M) \in \{1, \cdots, q_1\} \times \cdots \times \{1, \cdots, q_M\}$ we associate the global index

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

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 $n = 1, 2, \ldots, N$, the Lagrange basis $\{l_{n,j}\}_{j=1}^{q_n}$ of the space $P_{q_{n-1}}$,

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

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 $q - 1 = (q_1 - 1, \ldots, q_M - 1)$, that is, $P_{q-1}(\Gamma) = \bigotimes_{i=1}^{M} P_{q_i-1}(\Gamma_i)$. 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

$$E_q^{GL}[g] = \sum_{k=1}^{N_q} \omega_k g(y_k), \quad \omega_k = \prod_{n=1}^{M} \omega_{k_n}, \quad \omega_{k_n} = \int_{\Gamma_n} l_{k_n}^2(y)\,\mathrm{d}y\ . \tag{46}$$

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 state and adjoint problems to the complex domain. To do so, with slight abuse of notation let

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

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

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

We consider the state and adjoint problems extended to the complex domain: $\forall z \in \mathcal{U}_0$ find $y(\cdot, z) \in H_0^1(D; \mathbb{C})$ such that

$$\int_D a(x,z)\nabla y(x,z)\nabla v(x)\mathrm{d}x = \int_D (\widehat{u}(x) + g(x))v(x)\mathrm{d}x \quad \forall v \in H_0^1(D;\mathbb{C}) \,, \tag{47}$$

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

$$\int_D a(x,z)\nabla p(x,z)\nabla v(x)\mathrm{d}x = \int_D (y(x,z) - z_d(x))v(x)\mathrm{d}x \quad \forall v \in H_0^1(D;\mathbb{C}) \,. \tag{48}$$

It is well known that problem (47) and (48) 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 $a_{min} = \operatorname{essinf}_{x\in D} a_0(x)$. 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] for example, so that we only give the proof for $z \mapsto p(\cdot, z)$.

Lemma 9. *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*

$$\max_{z\in\Sigma} \|y(\cdot, z)\|_{H_0^1} \leq C_P \frac{\|g + \widehat{u}\|}{a_{min}} \tag{49}$$

and

$$\max_{z\in\Sigma} \|p(\cdot, z)\|_{H_0^1} \leq C_P \frac{\|z_d\|}{a_{min}} + C_P^3 \frac{\|g + \widehat{u}\|}{a_{min}^2} \,. \tag{50}$$

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 (49). This property translates to the adjoint 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 8. *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 exist $C > 0$ 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 (36) in Remark 2. In that case the factor c_j in (30) 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 $N_j \sim j^{\tau_0 l - 1}$ with $\tau_0 l - 1 > 0$. We use the recursive formula (30) and set, as before, $u_\star^h$ to be the exact optimal control for the FE problem defined in (11). We emphasize that (11) has no approximation in probability space. Setting $a_j = \mathbb{E}[\|u_j^h - u_\star^h\|^2]$ and $\beta_j = \frac{2\tau_j^2}{N_j}\mathbb{E}[\|\nabla f(u_\star^h, \omega)\|^2]$, we have from (30) applied to the sequence of FE solutions $\{u_j^h\}_{j>0}$ that

$$\begin{aligned} a_{j+1} \leq & c_j a_j + \beta_j \\ \leq & c_j c_{j-1} a_{j-1} + c_j \beta_{j-1} + \beta_j \\ \leq & \cdots \\ \leq & \underbrace{\Big(\prod_{i=1}^{j} c_i\Big)}_{=\kappa_j} a_1 + \underbrace{\sum_{i=1}^{j} \beta_i \prod_{l=i+1}^{j} c_l}_{=\mathcal{B}_j} . \end{aligned} \tag{51}$$

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 (51) we have

$$\mathcal{B}_j = \sum_{i=1}^{j} \beta_i \prod_{k=i+1}^{j} c_k \leq \sum_{i=1}^{j} S' i^{-\tau_0 l - 1} \underbrace{\prod_{k=i+1}^{j} \Big(1 - \frac{\tau_0 l}{k} + \frac{3\tau_0^2 L^2}{k^2}\Big)}_{=K_{ij}} ,$$

with $S' = 2\tau_0^2 \mathbb{E}[\|\nabla f(u_\star^h, \omega)\|^2]$. For the term K_{ij} we find that

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

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

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

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

$$a_{j+1} \leq D_3 j^{-\tau_0 l} a_1 + D_4 j^{-\tau_0 l} \log(j). \tag{52}$$

We conclude that

$$a_{j+1} \leq D_4 j^{-\tau_0 l} \log(j), \tag{53}$$

with D_4 possibly depending on $\|u_0^h - u_\star^h\|$. Finally, splitting the error as

$$\mathbb{E}[\|u_j^h - u_\star\|^2] \leq 2\mathbb{E}[\|u_j^h - u_\star^h\|^2] + 2\mathbb{E}[\|u_\star^h - u_\star\|^2]$$

and using (16) 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] A. Cotter, O. Shamir, N. Srebro, and K. Sridharan, *Better mini-batch algorithms via accelerated gradient methods*, Advances in Neural Information Processing Systems 24 (J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, eds.), Curran Associates, Inc., 2011, pp. 1647–1655.
[13] J.C. De los Reyes, *Numerical PDE-constrained optimization*, SpringerBriefs in Optimization, Springer, Cham, 2015. MR 3308473
[14] A. Défossez and F. Bach, *Averaged least-mean-squares: Bias-variance tradeoffs and optimal sampling distributions*, Artificial Intelligence and Statistics, 2015, pp. 205–213.
[15] O. Dekel, R. Gilad-Bachrach, O. Shamir, and L. Xiao, *Optimal distributed online prediction using mini-batches*, J. Mach. Learn. Res. **13** (2012), 165–202.
[16] A. Dieuleveut and F. Bach, *Nonparametric stochastic approximation with large step-sizes*, The Annals of Statistics **44** (2016), no. 4, 1363–1399.
[17] J. Duchi and Y. Singer, *Efficient online and batch learning using forward backward splitting*, J. Mach. Learn. Res. **10** (2009), 2899–2934.
[18] L. C. Evans, *Partial differential equations*, Graduate studies in mathematics, American Mathematical Society, 1998.
[19] N. Flammarion and F. Bach, *From averaging to acceleration, there is only a step-size*, Conference on Learning Theory, 2015, pp. 658–695.
[20] C. Geiersbach and W. Wollner, *A stochastic gradient method with mesh refinement for PDE constrained optimization under uncertainty*, Preprint 1905.08650, arXiv, 2019.
[21] P. Grisvard, *Elliptic problems in nonsmooth domains*, Classics in Applied Mathematics, vol. 69, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011, Reprint of the 1985 original [MR0775683], With a foreword by Susanne C. Brenner. MR 3396210
[22] 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
[23] 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.
[24] P. A. Guth, V. Kaarnioja, F. Y. Kuo, C. Schillings, and I. H. Sloan, *A quasi-Monte Carlo method for an optimal control problem under uncertainty*, arXiv:1910.10022, 2019.
[25] S. B. Hazra, *Large-scale PDE-constrained optimization in applications*, Springer-Verlag, Berlin, 2010.
[26] F. Hecht, *New development in freefem++*, J. Numer. Math. **20** (2012), no. 3-4, 251–265. MR 3043640

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

[28] 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

[29] 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

[30] 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

[31] ———, *Existence and optimality conditions for risk-averse PDE-constrained optimization*, SIAM/ASA J. Uncertain. Quantif. **6** (2018), no. 2, 787–815. MR 3810513

[32] 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

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

[34] 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.

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

[36] M. Martin, *Stochastic approximation methods for pde constrained optimal control problems with uncertain parameters*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 2019, Thesis n. 7233.

[37] 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.

[38] 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.

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

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

[41] 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

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

[43] 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).

[44] 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.
[45] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming*, second ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2014, Modeling and theory.
[46] 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
[47] A. Van Barel and S. Vandewalle, *Robust optimization of PDEs with random coefficients using a multilevel Monte Carlo method*, SIAM/ASA J. Uncertain. Quantif. **7** (2019), no. 1, 174–202. MR 3904437
[48] D. Williams, *Probability with martingales*, Cambridge mathematical textbooks, Cambridge University Press, 1991.