

## Mean-Variance Risk-Averse Optimal Control of Systems Governed by PDEs with Random Parameter Fields Using Quadratic Approximations\*

Alen Alexanderian<sup>†</sup>, Noemi Petra<sup>‡</sup>, Georg Stadler<sup>§</sup>, and Omar Ghattas<sup>¶</sup>

**Abstract.** We present a method for optimal control of systems governed by partial differential equations (PDEs) with uncertain parameter fields. We consider an objective function that involves the mean and variance of the control objective, leading to a risk-averse optimal control problem. Conventional numerical methods for optimization under uncertainty are prohibitive when applied to this problem. To make the optimal control problem tractable, we invoke a quadratic Taylor series approximation of the control objective with respect to the uncertain parameter field. This enables deriving explicit expressions for the mean and variance of the control objective in terms of its gradients and Hessians with respect to the uncertain parameter. The risk-averse optimal control problem is then formulated as a PDE-constrained optimization problem with constraints given by the forward and adjoint PDEs defining these gradients and Hessians. The expressions for the mean and variance of the control objective under the quadratic approximation involve the trace of the (preconditioned) Hessian and are thus prohibitive to evaluate. To overcome this difficulty, we employ trace estimators, which only require a modest number of Hessian-vector products. We illustrate our approach with two specific problems: the control of a semilinear elliptic PDE with an uncertain boundary source term, and the control of a linear elliptic PDE with an uncertain coefficient field. For the latter problem, we derive adjoint-based expressions for efficient computation of the gradient of the risk-averse objective with respect to the controls. Along with the quadratic approximation and trace estimation, this ensures that the cost of computing the risk-averse objective and its gradient with respect to the control—measured in the number of PDE solves—is independent of the (discretized) parameter and control dimensions, and depends only on the number of random vectors employed in the trace estimation, leading to an efficient quasi-Newton method for solving the optimal control problem. Finally, we present a comprehensive numerical study of an optimal control problem for fluid flow in a porous medium with an uncertain permeability field.

**Key words.** optimization under uncertainty, PDE-constrained optimization, optimal control, risk-aversion, PDEs with random coefficients, Gaussian measure, Hessian, trace estimators

**AMS subject classifications.** 60H15, 60H35, 35Q93, 35R60, 65K10

**DOI.** 10.1137/16M106306X

**1. Introduction.** An important class of problems arising in engineering and science is the optimization or optimal control of natural or engineered systems governed by partial

\*Received by the editors February 26, 2016; accepted for publication (in revised form) September 7, 2017; published electronically November 21, 2017.

<http://www.siam.org/journals/juq/5/M106306.html>

**Funding:** This work was partially supported by NSF grants 1508713 and 1507009, and by DOE grants DE-FC02-13ER26128, DE-SC0010518, and DE-FC02-11ER26052.

<sup>†</sup>Department of Mathematics, North Carolina State University, Raleigh, NC 27695 ([alexanderian@ncsu.edu](mailto:alexanderian@ncsu.edu)).

<sup>‡</sup>Applied Mathematics, School of Natural Sciences, University of California, Merced, CA 95340 ([npetra@ucmerced.edu](mailto:npetra@ucmerced.edu)).

<sup>§</sup>Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 ([stadler@cims.nyu.edu](mailto:stadler@cims.nyu.edu)).

<sup>¶</sup>Institute for Computational Engineering and Sciences, Department of Mechanical Engineering, and Department of Geological Sciences, The University of Texas at Austin, Austin, TX 78712 ([omar@ices.utexas.edu](mailto:omar@ices.utexas.edu)).

differential equations (PDEs). Often, the PDE models of these systems are characterized by parameters (or parameter functions) that are not known and are considered uncertain and hence modeled as random variables. These parameters can appear, for example, as coefficients, boundary data, initial conditions, or source terms. Consequently, optimization of such systems should be done in a way that the computed optimal controls or designs are robust with respect to the variability in the uncertain parameters.

*Literature survey and challenges.* There is a rich body of literature on theoretical and computational aspects of optimal control of systems governed by PDEs [45, 25, 5, 9, 27] and of optimization under uncertainty (OUU) [6, 40, 4, 41]. Recently there has been considerable interest in solution methods for optimization problems lying at the intersection of these two fields, namely, optimization problems governed by PDEs with uncertain parameters [10, 9, 11, 26, 28, 44, 29, 32, 31, 34, 30, 13, 33, 16]. To discuss the challenges of optimal control and, more generally, optimization of systems governed by PDEs with uncertain parameters, we consider a real-valued optimization objective  $\Theta(z, m)$  that depends on a control variable  $z$  and an uncertain parameter  $m$ , both of which can be finite- or infinite-dimensional. Throughout this article we refer to  $\Theta(z, m)$  as the control objective. The evaluation of this control objective requires the solution of a system of PDEs. Namely,  $\Theta(z, m) := \tilde{\Theta}(z, m, u)$  with  $u = \mathcal{S}(z, m)$ , where  $\mathcal{S}$  is a PDE solution operator. Here, we assumed that the system of PDEs admits a unique solution  $u$  for every pair  $(z, m)$  of controls and parameters. This dependence of  $\Theta$  on the solution of a PDE makes the evaluation (and the computation of derivatives) of  $\Theta$  computationally expensive. The presence of uncertain parameters greatly compounds the computational challenges of solving the PDE-constrained optimization problem.

In an OUU problem, it is natural to seek optimal controls  $z$  that make  $\Theta$  small in an average sense. For example, a risk-neutral optimal control approach seeks controls that solve

$$(1) \quad \min_z \mathbb{E}\{\Theta(z, m)\},$$

where  $\mathbb{E}\{\cdot\}$  denotes expectation over the uncertain parameter  $m$ . If we seek controls that, in addition to minimizing the expected value of  $\Theta$  with respect to  $m$ , result in a small uncertainty in  $\Theta$ , we are led to risk-averse optimal control. In the present work, we use the variance of the control objective as a risk measure and seek optimal controls that solve the problem

$$(2) \quad \min_z \mathbb{E}\{\Theta(z, m)\} + \gamma \text{Var}\{\Theta(z, m)\}.$$

Here,  $\text{Var}\{\cdot\}$  denotes the variance with respect to  $m$ , and  $\gamma > 0$  is a *risk-aversion* parameter that aims to penalize large variances of the control objective. This *mean-variance* formulation is only one of several formulations for finding risk-averse optimal controls. Other examples of more complex risk measures include the value at risk (VaR) and the conditional value at risk (CVaR) [37, 41]. Compared to the mean-variance formulation, these approaches do not symmetrically penalize the deviation of the quantity of interest around the mean, which is desirable, for instance, in applications where the control objective models a loss. In [33], the authors consider primal and dual formulations of a risk-averse PDE-constrained OUU problem using CVaR. They employ and study smooth approximations of the primal formulation to enable the application of derivative-based optimization methods to the CVaR objective, and rely on quadrature-based discretizations in the discretized parameter space.

To illustrate the main computational challenges involved in OUU problems, let us consider the (simpler) risk-neutral problem. A common approach to cope with the expectation in the objective function uses sampling over the random parameter space,  $E\{\Theta(z, m)\} \approx \sum_{i=1}^n w_i \Theta(z, m_i)$ , where  $\{m_i\}_{i=1}^n$  is a sample set, and  $w_i$  are sample weights. In the context of PDE-constrained OUU, evaluation of  $\Theta(z, m_i)$  requires solving the PDE problem  $u_i = \mathcal{S}(z, m_i)$  for each sample point  $m_i$ . The sample set  $\{m_i\}_{i=1}^n$  is chosen either by Monte Carlo sampling, where each  $m_i$  is a draw from the distribution law of  $m$  (and  $w_i = 1/n$  for every  $i$ ), or, for a suitably low-dimensional parameter space, is based on quadrature rules (and  $w_i$  are quadrature weights). The Monte Carlo-based approach, sometimes referred to as sample average approximation (SAA), is computationally prohibitive for OUU problems governed by PDEs. This is due to the slow convergence of Monte Carlo and the resulting large number of PDE solves for each evaluation of the expectation. Quadrature-based methods, obtained from tensorization of one-dimensional quadrature rules, use the regularity of  $\Theta(z, m)$  with respect to  $m$  and can accelerate convergence, but they suffer from the curse of dimensionality—the exponential growth of the number of quadrature points as the dimension increases. The use of sparse quadrature [42] can mitigate but not overcome the curse of dimensionality. Quadrature-based methods can be improved significantly by using adaptive sparse grids (see, e.g., [32, 9, 31], in which adaptive sparse grids are employed to solve OUU problems); however, these approaches are still computationally expensive for problems with parameter dimensions on the order of hundreds or thousands. Another class of methods for OUU problems are stochastic approximation (SA) methods [36, 24, 22, 39]. Similar to methods based on SAA, SA methods are computationally intractable for PDE-constrained OUU problems with high-dimensional parameters due to their slow convergence and the resulting need for a prohibitively large number of PDE solves.

*Approach.* We consider an uncertain parameter  $m$  that is modeled with a random field, which can also be viewed as a function-valued random variable. In the uncertainty quantification literature, it is common to use an a priori dimension reduction provided by a truncated Karhunen–Loève (KL) decomposition for such problems. However, KL modes that appear unimportant in simulating the random process may turn out to be important to the control objective. Moreover, a priori truncation of the KL expansion of a random field is most useful if the eigenvalues of the covariance operator (of the uncertain parameter) exhibit rapid decay. This is not always the case, for example, in the presence of small correlation lengths; in such cases, an a priori truncation needs to retain a large number of KL modes. We do not follow such approaches, to avoid bias introduced by the truncated KL expansion. Instead, we seek formulations that preserve the problem’s infinite-dimensional character and work in an infinite-dimensional setting as long as possible. Moreover, we aim to devise algorithms whose computational complexity, measured in the number of PDE solves, is independent of the discretized parameter dimension.

In the present work, we employ quadratic approximations of the parameter-to-objective map,  $m \mapsto \Theta(\cdot, m)$ , to render the computation of the control objective and its gradient (as necessitated by a gradient-based optimization method) tractable. Related approaches for OUU with finite-dimensional uncertain parameters and inexpensive-to-evaluate (compared to problems governed by PDEs) control objectives are used in [18, 17]. More generally, linear or quadratic expansions with respect to uncertain finite-dimensional parameters have also

been used for robust (finite-dimensional) optimization and reliability methods in engineering applications; we refer the reader to, e.g., [21, 35]. Using this approach, we can compute the moments of the first- and second-order Taylor expansions of  $\Theta(z, m)$  analytically. For optimal control problems with infinite-dimensional parameters, computation of the derivatives with respect to the uncertain parameters and the controls is prohibitive using the direct sensitivity approach (or finite differences). Instead, we employ adjoint methods to avoid dependence on the dimension of the discretized parameter field. Our formulation is particularized to two model problems: the control of a semilinear elliptic PDE with an uncertain boundary source term, and the control of a linear elliptic PDE with an uncertain coefficient field. The latter is motivated by industrial problems involving the optimal control of flows in porous media.

As we will see, using the quadratic approximation of  $m \mapsto \Theta(\cdot, m)$  results in an OOU objective function that involves traces of operators that depend on the Hessian of this mapping. Since direct computation of these traces is prohibitive for high-dimensional problems (explicit computation of the Hessian requires as many PDE solves as there are parameters), we use trace estimation based either on random vectors or on eigenvectors of the (preconditioned) Hessian at a nominal control. This only requires the action of the Hessian on vectors and is thus well suited for control problems governed by systems of PDEs.

**Contributions.** The main contributions of this work are as follows: (1) For an uncertain parameter field that follows a Gaussian distribution law, we derive analytic expressions for the mean and variance of a quadratic approximation to the parameter-to-objective map in infinite dimensions. These results are the basis for developing an efficient OOU approach that extends the work in [17] to a method suitable for large-scale PDE-constrained OOU problems. (2) We propose a formulation of the risk-averse OOU problem as a PDE-constrained optimization problem, with the constraints given by the PDEs defining the adjoint-based expression for the gradient  $\Theta_m$  and the linear action of the Hessian  $\Theta_{mm}$  of the parameter-to-objective map. Our method ensures that the cost of computing the risk-averse objective and its gradient with respect to the control—measured in the number of PDE solves—is independent of the (discretized) parameter and control dimensions, and depends only on the number of random vectors used in the trace estimation. (3) We fully elaborate our approach for the risk-averse control of an elliptic PDE with uncertain coefficient field and, in particular, derive the adjoint-based gradient of the control objective with respect to the control. We numerically study various aspects of our risk-aversion measure and of the efficiency of the method. The results show the effectiveness of our approach in computing risk-averse optimal controls in a problem with a 3,000-dimensional discretized parameter space.

**Limitations.** We also remark on limitations of our method. (1) Since we rely on approximations based on Taylor expansions, our arguments require smoothness (and proper boundedness of the derivatives) of the parameter-to-objective map. For the mean-variance formulation used here, this smoothness mainly depends on the governing PDEs and how the parameter enters these PDEs. (2) Compared to sampling-based methods, our approach requires first and second derivatives of  $\Theta(z, m)$  with respect to  $m$ , and thus appropriate adjoint solvers for the efficient computation of these derivatives. However, efficient methods for the solution of optimal control problems governed by PDEs will require adjoint-based derivatives with respect to the control; only minor modifications are needed to obtain derivatives with respect to the uncertain parameters. (3) Our derivation of the mean and variance of the quadratic approx-

imation to the parameter-to-objective map assumes that the parameter space is a Hilbert space, say  $\mathcal{H}$ , and that  $\Theta(z, \cdot)$  is defined and has the required derivatives in  $\mathcal{H}$ . As such, our framework does not apply to cases where the parameter space is a general Banach space. Even when we consider a Hilbert space  $\mathcal{H}$  of uncertain parameters,  $\Theta(z, m)$  may not be defined for every  $m$  in  $\mathcal{H}$  and have derivatives that are bounded in  $\mathcal{H}$ , in which case the expressions for the mean and variance could be used only formally to obtain an objective function for a risk-averse OUU problem. This is the case for the linear elliptic PDE problem with uncertain coefficient field, where the parameter-to-objective map is defined and has the required smoothness only in a subspace that has full measure. In section 3.3, we discuss such issues further and give conditions that ensure that the expressions for the mean and variance are well defined. Extensions to more general cases will be a subject of our future work.

**2. Preliminaries.** We let  $\mathcal{H}$  be an infinite-dimensional real separable Hilbert space endowed with an inner product  $\langle \cdot, \cdot \rangle$  and induced norm  $\|\cdot\|^2 = \langle \cdot, \cdot \rangle$ . We consider uncertain parameters that are modeled as spatially distributed random processes, which can be viewed as function-valued random variables. Let  $m$  denote such an uncertain parameter. We assume that the distribution law of  $m$ , which we denote by  $\mu$ , is supported on  $\mathcal{H}$  and consider  $m$  as an  $\mathcal{H}$ -valued random variable. That is,  $m$  is a function,  $m : (\Omega, \Sigma, P) \rightarrow (\mathcal{H}, \mathcal{B}(\mathcal{H}))$ , where  $\Omega$  is a sample space,  $\Sigma$  is an appropriate sigma-algebra, and  $P$  is a probability measure; here,  $\mathcal{B}(\mathcal{H})$  denotes the Borel sigma-algebra on  $\mathcal{H}$ . In what follows, with a slight abuse of notation, we denote the realizations of the uncertain parameter using the same symbol  $m$ .

As is common practice, instead of working on the abstract probability space  $(\Omega, \mathcal{F}, P)$ , we work in the image space  $(\mathcal{H}, \mathcal{B}(\mathcal{H}), \mu)$ , where  $\mu$  is the law of  $m$ , and use  $E\{\phi(m)\} = \int_{\mathcal{H}} \phi(m) \mu(dm)$  for an integrable function  $\phi : \mathcal{H} \rightarrow \mathbb{R}$ . As mentioned in the introduction, we assume that  $m$  has a Gaussian probability law,  $\mu = \mathcal{N}(\bar{m}, \mathcal{C})$ . Here,  $\mathcal{C}$  is a self-adjoint, positive trace class operator, and thus  $\mu$  defines a Gaussian measure on  $\mathcal{H}$ .

We denote by  $\mathcal{Z}_{\text{ad}}$  the set of admissible controls, which is a closed convex subset of  $L^2(\mathcal{D})$ , where  $\mathcal{D} \subset \mathbb{R}^d$  is a bounded open set with piecewise smooth boundary. We consider the control objective  $\Theta$  as a real-valued function defined on  $\mathcal{Z}_{\text{ad}} \times \mathcal{H}$ . We refer to the mapping  $m \mapsto \Theta(\cdot, m)$  as the parameter-to-objective map. For normed linear spaces  $\mathcal{X}$  and  $\mathcal{Y}$  we denote by  $\mathcal{L}(\mathcal{X}, \mathcal{Y})$  the space of bounded linear transformations from  $\mathcal{X}$  to  $\mathcal{Y}$ , and by  $\mathcal{L}(\mathcal{X})$  the space of bounded linear operators on  $\mathcal{X}$ . For a Hilbert space  $\mathcal{H}$ , we use  $\mathcal{L}_{\text{sym}}(\mathcal{H})$  to denote the subspace of  $\mathcal{L}(\mathcal{H})$  consisting of self-adjoint linear operators.

**2.1. Linear and quadratic expansions.** Recall that for a function  $f : \mathcal{X} \rightarrow \mathbb{R}$ , where  $\mathcal{X}$  is a Banach space, existence of first and second Fréchet derivatives at  $\bar{m} \in \mathcal{X}$  implies that

$$(3) \quad f(m) = f(\bar{m}) + f'(\bar{m})[m - \bar{m}] + \frac{1}{2} f''(\bar{m})[m - \bar{m}, m - \bar{m}] + o(\|m - \bar{m}\|_{\mathcal{X}}^2),$$

with  $f'(\bar{m}) \in \mathcal{X}^*$  and  $f''(\bar{m}) \in \mathcal{L}(\mathcal{X}, \mathcal{X}^*)$ . In the following, we consider the case where  $\mathcal{X}$  is a Hilbert space  $\mathcal{H}$ , and hence the derivatives admit Riesz representers in  $\mathcal{H}$ .

Next, we consider the control objective,  $\Theta : \mathcal{Z}_{\text{ad}} \times \mathcal{H} \rightarrow \mathbb{R}$ , and assume for an arbitrary control,  $z \in \mathcal{Z}_{\text{ad}}$ , existence of first and second Fréchet derivatives for the parameter-to-objective map  $m \mapsto \Theta(z, m)$  at  $\bar{m} \in \mathcal{H}$ . Consider the linear and quadratic approximations of

the parameter-to-objective map:

$$(4) \quad \Theta_{\text{lin}}(z, m) = \Theta(z, \bar{m}) + \langle \Theta_m(z, \bar{m}), m - \bar{m} \rangle,$$

$$(5) \quad \Theta_{\text{quad}}(z, m) = \Theta(z, \bar{m}) + \langle \Theta_m(z, \bar{m}), m - \bar{m} \rangle + \frac{1}{2} \langle \Theta_{mm}(z, \bar{m})(m - \bar{m}), m - \bar{m} \rangle.$$

Notice that for clarity, we denote first and second derivatives of  $\Theta$  with respect to  $m$  by  $\Theta_m$  and  $\Theta_{mm}$ , respectively. In this paper, we mainly use the quadratic approximation  $\Theta_{\text{quad}}$ .

Using these approximations rather than the exact parameter-to-objective map  $\Theta(z, m)$  enables one to analytically compute the moments appearing in the objective function of the (approximate) risk-averse OUU problem. This computation is facilitated by the fact that  $m$  has a Gaussian distribution law. Note also that the accuracy of such linear and quadratic approximations, considered in an average sense, can be related to the variance of the uncertain parameter. In section 3, where we derive analytic expressions for the mean and variance of the local quadratic approximation to  $\Theta(z, m)$ , we also describe how the variance of  $m$  is related to the expected value of the truncation error in the quadratic approximation.

**2.2. Probability measures on  $\mathcal{H}$ .** Here we recall basics regarding Borel probability measures, and Gaussian measures on infinite-dimensional Hilbert spaces. Let  $\mu$  be a Borel probability measure on  $\mathcal{H}$  with finite first and second moments. The mean  $\bar{a}$  of  $\mu$  is an element of  $\mathcal{H}$  such that

$$\int_{\mathcal{H}} \langle s, b \rangle \mu(ds) = \langle \bar{a}, b \rangle \quad \text{for all } b \in \mathcal{H}.$$

The covariance operator  $\mathcal{C}$  of  $\mu$  is a positive self-adjoint trace-class operator that satisfies

$$\int_{\mathcal{H}} \langle a, s - \bar{a} \rangle \langle b, s - \bar{a} \rangle \mu(ds) = \langle \mathcal{C}a, b \rangle \quad \text{for all } a, b \in \mathcal{H}.$$

It is straightforward to show that  $\int_{\mathcal{H}} \|s - \bar{a}\|^2 \mu(ds) = \text{Tr}(\mathcal{C})$ , where  $\text{Tr}(\mathcal{C})$  denotes the trace of the (positive self-adjoint) operator  $\mathcal{C}$ ; see, e.g., [14, p. 8]. Note also that

$$(6) \quad \int_{\mathcal{H}} \langle b, s - \bar{a} \rangle^2 \mu(ds) = \langle \mathcal{C}b, b \rangle.$$

One can also show (see, e.g., [1, Lemma 1]) that for a bounded linear operator  $\mathcal{K} : \mathcal{H} \rightarrow \mathcal{H}$ ,

$$(7) \quad \int_{\mathcal{H}} \langle \mathcal{K}(s - \bar{a}), s - \bar{a} \rangle \mu(ds) = \text{Tr}(\mathcal{K}\mathcal{C}) = \text{Tr}(\mathcal{C}^{1/2}\mathcal{K}\mathcal{C}^{1/2}).$$

Now consider the case where the measure  $\mu$  is a Gaussian,  $\mu = \mathcal{N}(\bar{a}, \mathcal{C})$ , where  $\mathcal{C}$  is a positive, self-adjoint, trace-class operator. The mean  $\bar{a}$  is assumed to belong to the Cameron–Martin space  $\mathcal{E} := \text{Im}(\mathcal{C}^{1/2})$ . The Cameron–Martin space is a dense subspace of  $\mathcal{H}$  and is a Hilbert space endowed with the inner product  $\langle \cdot, \cdot \rangle_{\mathcal{E}} := \langle \mathcal{C}^{-1/2} \cdot, \mathcal{C}^{-1/2} \cdot \rangle$  [15]. While the space  $\mathcal{E}$  is dense in  $\mathcal{H}$ , it is in some sense very “thin”; more precisely,  $\mu(\mathcal{E}) = 0$ .

We follow the construction in [12] and define a Gaussian measure  $\mu$  on  $\mathcal{H} = L^2(\mathcal{D})$ , where  $\mathcal{D}$  is a bounded domain with piecewise smooth boundary, as follows: define the covariance operator as the inverse of the square of a Laplacian-like operator,

$$(8) \quad \mathcal{C} = (-\kappa\Delta + \alpha I)^{-2} =: \mathcal{A}^{-2},$$



where  $\kappa, \alpha > 0$ , and the domain of  $\mathcal{A}$  is given by  $D(\mathcal{A}) = \{u \in H^2(\mathcal{D}) : \nabla u \cdot \mathbf{n} = 0 \text{ on } \partial\mathcal{D}\}$ . Here,  $H^2(\mathcal{D})$  is the Sobolev space of  $L^2(\mathcal{D})$  functions with square integrable first and second weak derivatives, and  $\mathbf{n}$  is the unit outward normal for the boundary  $\partial\mathcal{D}$ . This construction of the operator  $\mathcal{C}$  ensures that it is positive, self-adjoint, and of trace-class, and thus the Gaussian measure  $\mu = \mathcal{N}(\bar{a}, \mathcal{C})$  on  $\mathcal{H}$  is well-defined. A Gaussian random field whose law is given by such a Gaussian measure has almost surely continuous realizations [43].

As we saw in (7), given a Borel probability measure on  $\mathcal{H}$  with bounded first and second moments, we can obtain a simple expression for the first moment of a quadratic form on  $\mathcal{H}$ . If  $\mu$  is a Gaussian measure  $\mathcal{N}(\bar{a}, \mathcal{C})$  on  $\mathcal{H}$ , we can also compute the second moment of a quadratic form (see Remark 1.2.9 in [15]). In particular, if we let  $\mathcal{K}$  be a self-adjoint bounded linear operator on  $\mathcal{H}$ , then

$$(9) \quad \int_{\mathcal{H}} \langle \mathcal{K}(s - \bar{a}), s - \bar{a} \rangle^2 \mu(ds) = 2\text{Tr}[(\mathcal{C}^{1/2} \mathcal{K} \mathcal{C}^{1/2})^2] + \text{Tr}(\mathcal{C}^{1/2} \mathcal{K} \mathcal{C}^{1/2})^2.$$

### 3. Risk-averse OUU with quadratic approximation of the parameter-to-objective map.

As discussed above, we consider a control objective  $\Theta = \Theta(z, m)$ , where  $m$  has a Gaussian distribution law  $\mu = \mathcal{N}(\bar{m}, \mathcal{C})$ . In section 3.1, we analytically derive the moments of the quadratic approximation  $\Theta_{\text{quad}}$  of  $\Theta$ . We discuss the approximation errors due to this approximation by studying the expected value of the remainder term in the Taylor expansion. In section 3.2, using the expressions for the moments of the quadratic approximation, we formulate the optimization problem for finding risk-averse optimal controls. Extensions of our OUU approach to problems where  $\Theta(z, m)$  is defined only in a subspace of  $\mathcal{H}$  are discussed in section 3.3.

**3.1. Quadratic approximation to a function of a Gaussian random variable.** In this section, we compute mean and variance of  $\Theta_{\text{quad}}$  defined in (5) in the infinite-dimensional Hilbert space setting. The following arguments are pointwise in the control  $z \in \mathcal{Z}_{\text{ad}}$ , and hence, for notational convenience, we suppress the dependence of  $\Theta$  on  $z$ . We begin by establishing the following technical result.

**Lemma 3.1.** *Let  $\mu = \mathcal{N}(\bar{a}, \mathcal{C})$  be a Gaussian measure on  $\mathcal{H}$ , let  $b \in \mathcal{H}$  be fixed, and let  $\mathcal{K}$  be a bounded linear operator on  $\mathcal{H}$ . Then*

$$\int_{\mathcal{H}} \langle b, s - \bar{a} \rangle \langle \mathcal{K}(s - \bar{a}), s - \bar{a} \rangle \mu(ds) = 0.$$

*Proof.* Without loss of generality, we assume  $\bar{a} = 0$ . Let  $\{e_i\}_1^\infty$  be an orthonormal basis of eigenvectors of  $\mathcal{C}$  with corresponding positive eigenvalues  $\{\lambda_i\}_1^\infty$ . By  $\pi_n(s) = \sum_{j=1}^n \langle s, e_j \rangle e_j$  we denote the orthogonal projection onto the span of the first  $n$  eigenvectors. Observe that  $\langle b, s \rangle \langle \mathcal{K}s, s \rangle = \lim_{n \rightarrow \infty} \langle b, \pi_n(s) \rangle \langle \mathcal{K}\pi_n(s), \pi_n(s) \rangle$  and that  $|\langle b, \pi_n(s) \rangle \langle \mathcal{K}\pi_n(s), \pi_n(s) \rangle| \leq \|\mathcal{K}\| \|b\| \|s\|^3$ . Since  $\int_{\mathcal{H}} \|s\|^3 \mu(ds) < \infty$ , we can apply the Lebesgue dominated convergence theorem to obtain

$$\begin{aligned} \int_{\mathcal{H}} \langle b, s \rangle \langle \mathcal{K}s, s \rangle \mu(ds) &= \lim_{n \rightarrow \infty} \int_{\mathcal{H}} \langle b, \pi_n(s) \rangle \langle \mathcal{K}\pi_n(s), \pi_n(s) \rangle \mu(ds) \\ &= \lim_{n \rightarrow \infty} \sum_{i,j,k=1}^n \langle b, e_i \rangle \langle \mathcal{K}e_j, e_k \rangle \int_{\mathcal{H}} \langle s, e_i \rangle \langle s, e_j \rangle \langle s, e_k \rangle \mu(ds) = 0, \end{aligned}$$

where in the last step we used that the random  $n$ -vector  $\mathbf{Y} : \mathcal{H} \rightarrow \mathbb{R}^n$  defined by  $\mathbf{Y}(s) = (\langle s, e_1 \rangle, \langle s, e_2 \rangle, \dots, \langle s, e_n \rangle)$  is an  $n$ -variate Gaussian whose distribution law is  $\mu \circ \mathbf{Y}^{-1} = \mathcal{N}(\mathbf{0}, \text{diag}(\lambda_1, \dots, \lambda_n))$ . Note that we also use the result (see, e.g., [46]) that for a mean zero  $n$ -variate normal random vector  $\mathbf{Y}$ ,  $\mathbb{E}\{Y_{\alpha_1} Y_{\alpha_2} Y_{\alpha_3}\} = 0$  for  $\alpha_1, \alpha_2, \alpha_3 \in \{1, \dots, n\}$ . ■

**3.1.1. Mean and variance of the quadratic approximation.** Next, we derive expressions for the mean and variance of  $\Theta_{\text{quad}}$  in the infinite-dimensional Hilbert space setting.

**Proposition 3.2.** *Let  $\Theta : (\mathcal{H}, \mathcal{B}(\mathcal{H}), \mu) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  be a function that is twice differentiable at  $\bar{m} \in \mathcal{H}$ , with gradient  $\Theta_m(\bar{m}) \in \mathcal{H}$  and Hessian  $\mathcal{H}(\bar{m}) \in \mathcal{L}_{\text{sym}}(\mathcal{H})$ . Let  $\Theta_{\text{quad}} : (\mathcal{H}, \mathcal{B}(\mathcal{H}), \mu) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  be as defined in (5). Then*

$$(10) \quad \mathbb{E}\{\Theta_{\text{quad}}\} = \Theta(\bar{m}) + \frac{1}{2} \text{Tr}[\mathcal{C}^{1/2} \Theta_{mm}(\bar{m}) \mathcal{C}^{1/2}],$$

$$(11) \quad \text{Var}\{\Theta_{\text{quad}}\} = \langle \Theta_m(\bar{m}), \mathcal{C}[\Theta_m(\bar{m})] \rangle + \frac{1}{2} \text{Tr}[(\mathcal{C}^{1/2} \Theta_{mm}(\bar{m}) \mathcal{C}^{1/2})^2].$$

*Proof.* The first statement follows from

$$\begin{aligned} \mathbb{E}\{\Theta_{\text{quad}}\} &= \int_{\mathcal{H}} \Theta_{\text{quad}}(m) \mu(dm) = \Theta(\bar{m}) + \frac{1}{2} \int_{\mathcal{H}} \langle \Theta_{mm}(\bar{m})(m - \bar{m}), m - \bar{m} \rangle \mu(dm) \\ &= \Theta(\bar{m}) + \frac{1}{2} \text{Tr}[\Theta_{mm}(\bar{m}) \mathcal{C}] = \Theta(\bar{m}) + \frac{1}{2} \text{Tr}[\mathcal{C}^{1/2} \Theta_{mm}(\bar{m}) \mathcal{C}^{1/2}]. \end{aligned}$$

To derive the expression for the variance, first note that the variance of  $\Theta_{\text{quad}}(m)$  equals the variance of  $\Theta_{\text{quad}}(m) - \Theta(\bar{m})$ . Thus,

$$(12) \quad \text{Var}\{\Theta_{\text{quad}}\} = \mathbb{E}\{(\Theta_{\text{quad}}(m) - \Theta(\bar{m}))^2\} - \mathbb{E}\{\Theta_{\text{quad}}(m) - \Theta(\bar{m})\}^2.$$

The first term on the right-hand side is given by

$$\begin{aligned} \mathbb{E}\{(\Theta_{\text{quad}}(m) - \Theta(\bar{m}))^2\} &= \mathbb{E}\left\{\left(\langle \Theta_m(\bar{m}), m - \bar{m} \rangle + \frac{1}{2} \langle \Theta_{mm}(\bar{m})(m - \bar{m}), m - \bar{m} \rangle\right)^2\right\} \\ &= \mathbb{E}\left\{\langle \Theta_m(\bar{m}), m - \bar{m} \rangle^2\right\} + \frac{1}{4} \mathbb{E}\left\{\langle \Theta_{mm}(\bar{m})(m - \bar{m}), m - \bar{m} \rangle^2\right\} \\ &\quad + \mathbb{E}\left\{\langle \Theta_m(\bar{m}), m - \bar{m} \rangle \langle \Theta_{mm}(\bar{m})(m - \bar{m}), m - \bar{m} \rangle\right\} \\ &= \langle \Theta_m(\bar{m}), \mathcal{C}[\Theta_m(\bar{m})] \rangle + \frac{1}{4} \text{Tr}[\mathcal{C}^{1/2} \Theta_{mm}(\bar{m}) \mathcal{C}^{1/2}]^2 + \frac{1}{2} \text{Tr}[(\mathcal{C}^{1/2} \Theta_{mm}(\bar{m}) \mathcal{C}^{1/2})^2]. \end{aligned}$$

This, along with  $\mathbb{E}\{\Theta_{\text{quad}}(m) - \Theta(\bar{m})\} = \frac{1}{2} \text{Tr}(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2})$  and (12), finishes the proof. ■

Note that the expressions for the mean and variance of the linear approximation  $\Theta_{\text{lin}}$  defined in (4) consist of only the first terms in (10) and (11), respectively. We also point out an intuitive interpretation of the *covariance-preconditioned* Hessian,  $\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2}$ , in the expressions for the mean and variance in (10) and (11). As the Hessian  $\Theta_{mm}$  only appears preconditioned by the covariance  $\mathcal{C}$ , the second-order contributions to the expectation and the variance are large only if the dominating eigenvector directions of  $\mathcal{C}$  and  $\Theta_{mm}$  are “similar.” More precisely, the eigenvectors of  $\mathcal{C}$  that correspond to large eigenvalues (i.e., directions of large uncertainty) only have a significant influence on the mean and variance if these directions



are also important for the Hessian of the prediction. Conversely, important directions for the prediction Hessian only result in significant contributions to the second-order approximation of mean and variance if the uncertainty in these directions is significant.

We point out that while the Gaussian assumption on the distribution law of  $m$  is required for derivation of the expression for  $\text{Var}\{\Theta_{\text{quad}}\}$  in Proposition 3.2, the expression for  $\text{E}\{\Theta_{\text{quad}}\}$  can be derived without this assumption. Namely, it holds if the law of  $m$  is any Borel probability measure on  $\mathcal{H}$  with bounded first and second moments. These assumptions on the law of  $m$  are also sufficient for deriving the expressions for the mean and variance of  $\Theta_{\text{lin}}$ ,

$$\text{E}\{\Theta_{\text{lin}}\} = \Theta(\bar{m}), \quad \text{Var}\{\Theta_{\text{lin}}\} = \langle \Theta_m(\bar{m}), \mathcal{C}[\Theta_m(\bar{m})] \rangle.$$

Here, the expression for the mean is immediate, and the one for variance follows from (6).

**3.1.2. Expected value of the truncation error.** Next, we discuss the error due to replacing  $\Theta$  by the quadratic approximation  $\Theta_{\text{quad}}$ . Assuming sufficient smoothness and boundedness of the derivatives of  $\Theta$ , we study the expected value of the truncation error as the uncertainty in the parameter  $m$  decreases; i.e., we consider  $\mathcal{N}(\bar{m}, \varepsilon\mathcal{C})$ , where  $\varepsilon > 0$  approaches zero. To gain intuition, consider the case when the covariance operator is such that parameter draws  $m$  have a high probability of being close to the mean  $\bar{m}$ , where the quadratic approximation is more accurate. In this case, we can expect the truncation error to be small.

To derive a quantitative estimate, we assume that  $\Theta$  is three times continuously differentiable such that the remainder term in the quadratic expansion (3) has the form

$$(13) \quad R(m; \bar{m}) := \Theta(m) - \Theta_{\text{quad}}(m) = \frac{1}{3!} \Theta^{(3)}(\xi)[m - \bar{m}, m - \bar{m}, m - \bar{m}].$$

Here,  $\Theta^{(3)}$  denotes the third derivative with respect to  $m$  at  $\xi$ , which is an element of the line segment between  $\bar{m}$  and  $m$ . Note that this and the following considerations are pointwise in the control variable  $z$ . We are interested in the expectation value of the remainder term, which we will relate to  $\text{Tr}(\mathcal{C})$ , the average variance of  $m$ .

Assuming  $\Theta^{(3)}$  is a uniformly bounded trilinear map, i.e., for all  $\xi \in \mathcal{H}$ ,  $|\Theta^{(3)}(\xi)[u, v, w]| \leq K \|u\| \|v\| \|w\|$  for all  $u, v, w \in \mathcal{H}$ , we obtain

$$(14) \quad \begin{aligned} \int_{\mathcal{H}} |R(m; \bar{m})| \mu(dm) &\leq K \int_{\mathcal{H}} \|m - \bar{m}\|^3 \mu(dm) = K \int_{\mathcal{H}} \|m - \bar{m}\| \|m - \bar{m}\|^2 \mu(dm) \\ &\leq K \left( \int_{\mathcal{H}} \|m - \bar{m}\|^2 \mu(dm) \right)^{1/2} \left( \int_{\mathcal{H}} \|m - \bar{m}\|^4 \mu(dm) \right)^{1/2}. \end{aligned}$$

Next, recall that  $\int_{\mathcal{H}} \|m - \bar{m}\|^2 \mu(dm) = \text{Tr}(\mathcal{C})$ . Moreover, using (9) with  $\mathcal{K} = I$ , we have

$$(15) \quad \int_{\mathcal{H}} \|m - \bar{m}\|^4 \mu(dm) = 2\text{Tr}(\mathcal{C}^2) + \text{Tr}(\mathcal{C})^2 \leq 3\text{Tr}(\mathcal{C})^2,$$

where we have also used  $\text{Tr}(\mathcal{C}^2) \leq \text{Tr}(\mathcal{C})^2$ . Therefore, using (14) and (15), we have

$$\int_{\mathcal{H}} |R(m; \bar{m})| \mu(dm) \leq \sqrt{3} K \text{Tr}(\mathcal{C})^{3/2}.$$

Now, if we consider a family of laws  $\mu_\varepsilon = \mathcal{N}(\bar{m}, \varepsilon\mathcal{C})$ ,  $\varepsilon > 0$ , for  $m$ , then the expected value of the remainder (13) is  $\mathcal{O}(\varepsilon^{3/2})$ , as  $\varepsilon \rightarrow 0$ . This should be contrasted with the expected value of the remainder term for the linear expansion, which can be shown to be  $\mathcal{O}(\varepsilon)$ . Note that if  $\Theta$  is cubic, the third derivative  $\Theta^{(3)}$  is constant, and the expectation over the remainder term vanishes since  $m$  follows a Gaussian distribution that is symmetric with respect to its mean.

The above argument regarding the expected value of the remainder in a Taylor expansion can be generalized for higher order expansions; see Appendix A.

**3.2. The OUU objective function.** We can now give the explicit form for the objective function for the risk-averse OUU problem (2), in which we use  $\Theta_{\text{quad}}$  rather than  $\Theta$ ; using (10) and (11), we consider the following expression for the OUU objective function:

$$(16) \quad \begin{aligned} \Theta(z, \bar{m}) + \frac{1}{2} \text{Tr}[\mathcal{C}^{1/2} \Theta_{mm}(z, \bar{m}) \mathcal{C}^{1/2}] \\ + \frac{1}{2} \left\{ \langle \Theta_m(z, \bar{m}), \mathcal{C}[\Theta_m(z, \bar{m})] \rangle + \frac{1}{2} \text{Tr}[(\mathcal{C}^{1/2} \Theta_{mm}(z, \bar{m}) \mathcal{C}^{1/2})^2] \right\} + \frac{\gamma}{2} \|z\|^2. \end{aligned}$$

Note that we have also added the control cost  $\frac{\gamma}{2} \|z\|^2$  in (16). The numerical computation of operator traces appearing in the expressions for the mean and variance of the quadratic approximation is expensive and can be prohibitive for inverse problems governed by PDEs. Hence, we employ approximations obtained by randomized trace estimators [3, 38], which require only the application of the operator to (random) vectors and provide reasonably accurate trace estimates using a small number  $n_{\text{tr}}$  of trace estimator vectors (see also [2, Appendix A] for a result on an infinite-dimensional Gaussian trace estimator). Trace estimation for  $\text{Tr}(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2})$  and  $\text{Tr}[(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2})^2]$  amounts to

$$(17) \quad \begin{aligned} \text{Tr}(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2}) &\approx \frac{1}{n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, \Theta_{mm} \zeta_j \rangle, \\ \text{Tr}[(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2})^2] &\approx \frac{1}{n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \Theta_{mm} \zeta_j, \mathcal{C}[\Theta_{mm} \zeta_j] \rangle, \end{aligned}$$

where  $\zeta_j$ ,  $j \in \{1, \dots, n_{\text{tr}}\}$  are draws from the measure  $\nu = \mathcal{N}(0, \mathcal{C})$ . This form of the trace estimators is justified by the identities

$$\int_{\mathcal{H}} \langle \Theta_{mm} \zeta, \zeta \rangle \nu(d\zeta) = \text{Tr}(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2}), \quad \int_{\mathcal{H}} \langle \Theta_{mm} \zeta, \mathcal{C}[\Theta_{mm} \zeta] \rangle \nu(d\zeta) = \text{Tr}[(\mathcal{C}^{1/2} \Theta_{mm} \mathcal{C}^{1/2})^2].$$

Replacing the operator traces in the OUU objective function (16) using trace estimators results in the OUU objective function

$$(18a) \quad \begin{aligned} \mathcal{J}(z) := \Theta(z, \bar{m}) + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, \psi_j \rangle \\ + \frac{1}{2} \left\{ \langle \Theta_m(z, \bar{m}), \mathcal{C}[\Theta_m(z, \bar{m})] \rangle + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \psi_j, \mathcal{C}\psi_j \rangle \right\} + \frac{\gamma}{2} \|z\|^2, \end{aligned}$$

where for  $j \in \{1, \dots, n_{\text{tr}}\}$ ,

$$(18b) \quad \psi_j = \Theta_{mm}(z, \bar{m})\zeta_j, \quad \zeta_j \sim \mathcal{N}(0, \mathcal{C}).$$

As an alternative to the randomized estimator in (17), we can use

$$\text{Tr}(\mathcal{C}^{1/2}\Theta_{mm}\mathcal{C}^{1/2}) \approx \sum_{j=1}^{n_{\text{tr}}} \langle w_j, \Theta_{mm}w_j \rangle, \quad \text{Tr}[(\mathcal{C}^{1/2}\Theta_{mm}\mathcal{C}^{1/2})^2] \approx \sum_{j=1}^{n_{\text{tr}}} \langle \Theta_{mm}w_j, \mathcal{C}[\Theta_{mm}w_j] \rangle,$$

where  $w_j = \mathcal{C}^{1/2}v_j$ ,  $\{v_j\}_{j=0}^{\infty}$  is an orthonormal basis of  $\mathcal{H}$ , and  $n_{\text{tr}}$  is an appropriate truncation level. One possibility is to choose  $v_j$ ,  $j \in \{1, \dots, n_{\text{tr}}\}$  as the dominant eigenvectors of  $\mathcal{C}^{1/2}\Theta_{mm}(\bar{m}, z_0)\mathcal{C}^{1/2}$ , where  $z_0$  is a nominal control variable. Since in many applications the operator  $\mathcal{C}^{1/2}\Theta_{mm}\mathcal{C}^{1/2}$  has a rapidly decaying spectrum,  $n_{\text{tr}}$  can be chosen small. While such an estimator is tailored to the Hessian evaluated at  $z_0$ , we have observed it to perform well for values of the control variable in a neighborhood of  $z_0$ . We will demonstrate the utility of this approach in our computational results.

Note that the existence of minimizers for  $\mathcal{J}(z)$  (as well as (16)) depends on the control space and on properties of  $\Theta_m$  and  $\Theta_{mm}$ , and must be argued on a case-by-case basis.

**3.3. Extensions.** In some applications, the control objective  $\Theta(z, m)$  might be defined only for  $m$  in a Banach subspace  $\mathcal{X} \subset \mathcal{H}$  with  $\mu(\mathcal{X}) = 1$ . Let  $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$  be such a subspace, and recall that since the Cameron–Martin space  $\mathcal{E}$  is compactly embedded in all subspaces of  $\mathcal{H}$  that have full measure [43],  $\mathcal{E}$  is compactly embedded in  $\mathcal{X}$ . Moreover, when using quadratic approximations, we need derivatives at  $\bar{m} \in \mathcal{E}$ . It is thus reasonable to require existence of derivatives only in the Cameron–Martin space. In such cases, one might be tempted to consider the restriction of  $\Theta$  to the Cameron–Martin space  $\mathcal{E}$  and use

$$(19) \quad \Theta(m) \approx \Theta(\bar{m}) + \langle \Theta_m(\bar{m}), m - \bar{m} \rangle_{\mathcal{E}, \mathcal{E}^*} + \frac{1}{2} \langle \Theta_{mm}(\bar{m})(m - \bar{m}), m - \bar{m} \rangle_{\mathcal{E}, \mathcal{E}^*}, \quad m \in \mathcal{E}.$$

Here  $\langle \cdot, \cdot \rangle_{\mathcal{E}, \mathcal{E}^*}$  denotes the duality pairing between  $\mathcal{E}$  and its dual  $\mathcal{E}^*$ , and  $\Theta_m(\bar{m}) \in \mathcal{E}^*$  and  $\Theta_{mm}(\bar{m}) \in \mathcal{L}(\mathcal{E}, \mathcal{E}^*)$  are the gradient and Hessian of  $\Theta(m)$  at  $m = \bar{m}$ , respectively. Note that we have suppressed the dependence of  $\Theta$  on  $z$ .

The definition (19), however, is not meaningful from a measure-theoretic point of view, as  $\mathcal{E}$  has measure zero. A possible remedy is to define a bounded and self-adjoint linear operator  $\mathcal{S}^\delta : \mathcal{H} \rightarrow \mathcal{E}$  and to consider the composition  $\Theta^\delta(m) := \Theta(\mathcal{S}^\delta m)$ . One possibility is to choose  $\mathcal{S}^\delta := (I + \delta \mathcal{C}^{-1/2})^{-1}$ , in which case the smoothing is controlled by  $\delta > 0$ . The gradient and Hessian of  $\Theta^\delta(\cdot)$  are now given by

$$\Theta_m^\delta(\bar{m}) = \mathcal{S}^\delta \Theta_m(\mathcal{S}^\delta \bar{m}) \in \mathcal{H}, \quad \Theta_{mm}^\delta(m) = \mathcal{S}^\delta \Theta_{mm}(\mathcal{S}^\delta \bar{m}) \mathcal{S}^\delta \in \mathcal{L}(\mathcal{H}).$$

This way, one might consider the local quadratic approximation

$$\Theta_{\text{quad}}^\delta(m) = \Theta^\delta(\bar{m}) + \langle \Theta_m^\delta(\bar{m}), m - \bar{m} \rangle + \frac{1}{2} \langle \Theta_{mm}^\delta(\bar{m})(m - \bar{m}), m - \bar{m} \rangle, \quad m \in \mathcal{H}.$$

This construction allows us to consider control objectives that are defined only on  $\mathcal{E}$  to be extended to  $\mathcal{H}$  via the mapping  $\Theta^\delta$ . Then the Hilbert space formulation of the risk-averse OUU with quadratic approximations, developed in earlier sections, can be applied.

Another case where the Hilbert space theory needs extension is when  $\Theta$  is defined on a Banach subspace  $\mathcal{X}$  of full measure, and thus its gradient and Hessian belong to  $\mathcal{X}^*$  and  $\mathcal{L}(\mathcal{X}, \mathcal{X}^*)$ , respectively. For example, in the control problem governed by a linear elliptic PDE with an uncertain coefficient, discussed in section 5, the space  $\mathcal{X} = C(\mathcal{D})$  plays such a role (it is known [43, 19] that due to our choice of the covariance operator  $\mathcal{C}$ ,  $\mu(\mathcal{X}) = 1$ ). In this case, we show that the expressions for the mean and variance of the quadratic approximation continue—under appropriate assumptions—to be well-defined.

Now, the gradient  $\Theta_m(\bar{m}) \in \mathcal{X}^* \subset \mathcal{E}^*$ , and thus  $\langle \Theta_m(\bar{m}), \mathcal{C}[\Theta_m(\bar{m})] \rangle$ , can be interpreted as a duality product, i.e., the linear action of  $\Theta_m(\bar{m}) \in \mathcal{E}^*$  on  $\mathcal{C}[\Theta_m(\bar{m})] \in \mathcal{E}$ . Here we have used that for the covariance operator  $\mathcal{C}$  defined above, we have  $\mathcal{C}^{1/2} : \mathcal{E}^* \rightarrow \mathcal{E}$ .

Next, considering the expressions (10) and (11) for mean and variance, it remains to specify conditions that ensure that the operator  $\mathcal{C}^{1/2}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2}$  is trace-class on  $\mathcal{H}$ .

**Proposition 3.3.** *Let  $\mathcal{C}$  be the covariance operator as defined in section 2.2. Assume that  $\Theta_{mm}(\bar{m}) \in \mathcal{L}(\mathcal{X}, \mathcal{X}^*)$  restricted to  $\mathcal{E}$  is a bounded linear operator on  $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$ . Then the operator  $\mathcal{C}^{1/2}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2}$  is a trace-class operator on  $\mathcal{H}$ .*

*Proof.* It is straightforward to see that  $\mathcal{C}^{1/2}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2} \in \mathcal{L}(\mathcal{H})$ . It remains to show that  $\mathcal{C}^{1/2}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2}$  is trace-class. Let  $\{e_j\}_{j=1}^{\infty}$  be the complete orthonormal set of eigenvectors of  $\mathcal{C}$ , with corresponding (positive) eigenvalues  $\{\lambda_j\}_{j=1}^{\infty}$ . We note that for  $u, v \in \mathcal{H}$ ,  $\langle u, v \rangle = \langle \mathcal{C}^{1/2}u, \mathcal{C}^{1/2}v \rangle_{\mathcal{E}}$ . Therefore, for each  $j \geq 1$ , we can write

$$\begin{aligned} \langle e_j, \mathcal{C}^{1/2}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2}e_j \rangle &= \langle \mathcal{C}^{1/2}e_j, \mathcal{C}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2}e_j \rangle_{\mathcal{E}} \leq \|\Theta_{mm}(\bar{m})\| \|\mathcal{C}^{1/2}\mathcal{C}e_j\|_{\mathcal{E}} \|\mathcal{C}^{1/2}e_j\|_{\mathcal{E}} \\ &= \|\Theta_{mm}(\bar{m})\| \|\mathcal{C}e_j\| \|e_j\| = \|\Theta_{mm}(\bar{m})\| \lambda_j. \end{aligned}$$

Therefore,  $\sum_j \langle e_j, \mathcal{C}^{1/2}\Theta_{mm}(\bar{m})\mathcal{C}^{1/2}e_j \rangle \leq \|\Theta_{mm}(\bar{m})\| \sum_j \lambda_j < \infty$ . ■

**4. Control of a semilinear elliptic PDE with uncertain Neumann boundary data.** We first illustrate our approach for the optimal control of a semilinear elliptic PDE with uncertain Neumann boundary data and a right-hand side control. In this problem, the nonlinearity in the governing PDE is the sole reason why the quadratic approximation  $\Theta_{\text{quad}}(\cdot)$  of the objective is not exact. Below, we present and discuss the optimization formulation for this PDE-constrained OUU problem.

We assume a bounded domain  $\mathcal{D} \subset \mathbb{R}^2$  with boundary split into disjoint parts  $\Gamma_N$  and  $\Gamma_D$ , and we consider the semilinear elliptic equation

$$(20) \quad \begin{aligned} -\Delta u + cu^3 &= z && \text{in } \mathcal{D}, \\ u &= 0 && \text{on } \Gamma_D, \\ \nabla u \cdot \mathbf{n} &= m && \text{on } \Gamma_N. \end{aligned}$$

Here,  $c \geq 0$ ,  $z \in L_2(\mathcal{D})$  is the control, and the uncertain parameter is  $m \in L^2(\Gamma_N)$ , distributed according to the law  $\mu = \mathcal{N}(\bar{m}, \mathcal{C})$  with mean  $\bar{m}$  and covariance operator  $\mathcal{C}$ . Due to the monotonicity of the nonlinear term, the state equation (20) has a unique solution for every Neumann data  $m$  and every  $z \in L_2(\mathcal{D})$  [20].

We consider a control objective of tracking type as follows:

$$(21) \quad \Theta(z, m) = \frac{1}{2} \|u - u_d\|^2,$$

where  $u_d \in L^2(\mathcal{D})$  is a given desired state. It is straightforward to show that the solution  $u$  of (20) satisfies the estimate  $\|u\|_{H^1(\mathcal{D})} \leq K(\|z\|_{L^2(\mathcal{D})} + \|m\|_{L^2(\Gamma_N)})$  for a constant  $K = K(\mathcal{D})$ . Hence, since  $m$  has moments of all orders, it follows that the state variable  $u = u(z, m)$  also has moments of all orders for every  $z \in L^2(\mathcal{D})$ . This in particular implies the existence of the first and second moments of the control objective  $\Theta(z, m)$  for every  $z$ . Therefore, a mean-variance risk-averse OUU objective function is well-defined.

To derive the quadratic approximation of (21) with respect to  $m$  at the mean  $\bar{m}$ , we compute, for fixed control  $z$ , the gradient and Hessian of  $\Theta$  with respect to  $m$ . The gradient of  $\Theta$  with respect to  $m$  is  $\Theta_m(\bar{m}) = -\bar{p}|_{\Gamma_N}$ , where  $u$  satisfies (20) with  $m = \bar{m}$  (the corresponding state is denoted by  $\bar{u}$ ), and  $\bar{p}$  satisfies the adjoint equation

$$(22) \quad \begin{aligned} -\Delta \bar{p} + 3c\bar{u}^2 \bar{p} &= -(\bar{u} - u_d) && \text{in } \mathcal{D}, \\ \bar{p} &= 0 && \text{on } \Gamma_D, \\ \nabla \bar{p} \cdot \mathbf{n} &= 0 && \text{on } \Gamma_N. \end{aligned}$$

The second derivative at  $\bar{m}$  evaluated in a direction  $\hat{m}$  is given by  $\Theta_{mm}(\bar{m})(\hat{m}) = -\hat{p}|_{\Gamma_N}$ , where  $\hat{p}$  solves the incremental adjoint equation

$$(23) \quad \begin{aligned} -\Delta \hat{p} + 3c\bar{u}^2 \hat{p} &= -(6c\bar{u}\bar{p} + 1)\hat{u} && \text{in } \mathcal{D}, \\ \hat{p} &= 0 && \text{on } \Gamma_D, \\ \nabla \hat{p} \cdot \mathbf{n} &= 0 && \text{on } \Gamma_N, \end{aligned}$$

and  $\hat{u}$  the incremental state equation

$$(24) \quad \begin{aligned} -\Delta \hat{u} + 3c\bar{u}^2 \hat{u} &= 0 && \text{in } \mathcal{D}, \\ \hat{u} &= 0 && \text{on } \Gamma_D, \\ \nabla \hat{u} \cdot \mathbf{n} &= \hat{m} && \text{on } \Gamma_N. \end{aligned}$$

We now show that  $\Theta_{mm}(\bar{m}) : m \mapsto -\hat{p}|_{\Gamma_N}$  is bounded as a mapping from  $L_2(\Gamma_N) \rightarrow L_2(\Gamma_N)$ . From (24) it follows that  $\|\hat{u}\|_{H^1(\mathcal{D})} \leq \|\hat{m}\|$ . To estimate the  $H^{-1}(\mathcal{D})$ -norm of the right-hand side in (23), we consider an arbitrary  $v \in H^1(\mathcal{D})$ . Using Hölder's inequality and the continuous embedding of  $H^1(\mathcal{D})$  in  $L^4(\mathcal{D})$ , we obtain

$$\begin{aligned} \int_{\mathcal{D}} (6c\bar{u}\bar{p} + 1)\hat{u}v \, dx &\leq \|6c\bar{u}\bar{p} + 1\| \|\hat{u}v\| \leq (c_1 + c_2\|\bar{u}\|_{L^4(\mathcal{D})}\|\bar{p}\|_{L^4(\mathcal{D})}) \|\hat{u}\|_{L^4(\mathcal{D})} \|v\|_{L^4(\mathcal{D})} \\ &\leq c_3 \|\hat{u}\|_{H^1(\mathcal{D})} \|v\|_{H^1(\mathcal{D})}, \end{aligned}$$

where the constants  $c_1, c_2, c_3$  do not depend on  $\hat{u}$  or  $v$ . This shows that the  $H^{-1}(\mathcal{D})$ -norm of the right-hand side in (24) is bounded by  $\|\hat{u}\|_{H^1(\mathcal{D})}$ . Hence, with constants  $c_4, c_5, c_6$ ,

$$\|\hat{p}|_{\Gamma_N}\|_{L^2(\Gamma_N)} \leq c_4 \|\hat{p}\|_{H^1(\mathcal{D})} \leq c_5 \|\hat{u}\|_{H^1(\mathcal{D})} \leq c_6 \|\hat{m}\|_{L^2(\Gamma_N)},$$

which proves the boundedness of  $\Theta_{mm}$ . Thus, the OUU objective function (16), specialized to the present example, with the gradient and Hessian operators  $\Theta_m(\bar{m})$  and  $\Theta_{mm}(\bar{m})$  defined above, is well-defined and conforms to the theory outlined in sections 3.1–3.2.

**5. Control of an elliptic PDE with an uncertain coefficient.** Motivated by problems involving the optimal control of flows in porous media, we consider the optimal control of a linear elliptic PDE with an uncertain coefficient field. We discuss this application numerically in section 6, where we consider control of fluid injection into the subsurface at injection wells. In this section, we describe the control objective and the PDE-constrained objective function for the risk-averse OUU problem, and derive the adjoint-based expressions for the gradient of the OUU objective function. This is an example of a problem in which the parameter-to-objective map is defined only on a Banach subspace of  $\mathcal{H}$  that has full measure. We thus use the expressions for the mean and variance of the quadratic approximation, developed in a Hilbert space setting in section 3, formally, to define the objective function for the risk-averse optimal control problem.

We begin by describing the state (forward) equation. On an open bounded and sufficiently smooth domain  $\mathcal{D} \subset \mathbb{R}^n$ ,  $n \in \{1, 2, 3\}$  with boundary  $\partial\mathcal{D}$ , we consider the following elliptic partial differential equation:

$$(25) \quad \begin{aligned} -\nabla \cdot (e^m \nabla u) &= b + Fz && \text{in } \mathcal{D}, \\ u &= g && \text{on } \Gamma_D, \\ e^m \nabla u \cdot \mathbf{n} &= 0 && \text{on } \Gamma_N. \end{aligned}$$

Here, the boundary is split into disjoint parts  $\Gamma_D$  and  $\Gamma_N$  on which we impose Dirichlet and Neumann boundary conditions, respectively. The Dirichlet data is  $g \in H^{1/2}(\Gamma_D)$ ,  $\mathbf{n}$  denotes the unit-length outward normal for the boundary  $\partial\mathcal{D}$ , and for simplicity we have considered homogeneous Neumann conditions. We assume that the right-hand side is specified as  $Fz + b$ , where  $F : L^2(\mathcal{D}) \rightarrow L^2(\mathcal{D})$  is a bounded linear transformation and  $b \in L^2(\mathcal{D})$  is a distributed source per unit volume.

We consider the weak form of (25), i.e., we seek solutions  $u \in \mathcal{V}_g := \{v \in H^1(\mathcal{D}) : v|_{\Gamma_D} = g\}$  that satisfy

$$(26) \quad \langle e^m \nabla u, \nabla v \rangle - \langle b + Fz, v \rangle = 0 \quad \text{for all } v \in \mathcal{V} := \{v \in H^1(\mathcal{D}) : v|_{\Gamma_D} = 0\}.$$

We consider the case where the log-permeability,  $m$ , is uncertain and is modeled as a spatially distributed random field; see also [19, 7]. The realizations of  $m$  belong to the Hilbert space  $\mathcal{H} = L^2(\mathcal{D})$ , and we assume  $m$  is distributed according to a Gaussian with covariance operator  $\mathcal{C}$  and mean  $\bar{m} \in \mathcal{E} \subset \mathcal{H}$ , where  $\mathcal{E}$  is the Cameron–Martin space associated with the Gaussian measure  $\mu$ . With our choice of the covariance operator,  $\mu(\mathcal{X}) = 1$  with  $\mathcal{X} = C(\bar{\mathcal{D}}) \subset \mathcal{H}$ .

We consider the following tracking-type control objective  $\Theta(z, m)$ :

$$(27) \quad \Theta(z, m) = \frac{1}{2} \|\mathcal{Q}u(z, m) - \bar{\mathbf{q}}\|_2^2,$$

where  $u = u(z, m)$  solves the weak form of the state equation (26),  $\mathcal{Q} : L^2(\mathcal{D}) \rightarrow \mathbb{R}^q$  is a bounded linear operator, and  $\bar{\mathbf{q}} \in \mathbb{R}^q$  is given. Notice that this control objective is defined for every  $m \in \mathcal{X}$ , i.e., almost surely. It is also possible to prove boundedness of moments of  $\|u\|_{\mathcal{X}}$ , which, in particular, ensures that a mean-variance risk-averse OUU objective based on (27) is well-defined. Compared to the problem discussed in section 4, here the proof of



boundedness of moments of  $\|u\|_{\mathcal{X}}$  is more involved (see [19, Example 2.15]) and requires the use of Fernique's theorem [23].

For fixed parameter  $m$  and control  $z$ , the gradient of the parameter-to-objective map can be computed with a standard variational calculus approach (see, e.g., [8]). In particular, at  $m = \bar{m}$ , the gradient  $\Theta_m(\bar{m}) \in \mathcal{X}^*$  is given by

$$(28) \quad \Theta_m(\bar{m}) = e^{\bar{m}} \nabla u \cdot \nabla p,$$

where  $u$  is the solution to the state equation, and  $p$  solves the adjoint equation, i.e.,  $p \in \mathcal{V}$ , and satisfies

$$(29) \quad \langle e^{\bar{m}} \nabla p, \nabla \tilde{p} \rangle = -\langle \mathcal{Q}^*(\mathcal{Q}u - \bar{\mathbf{q}}), \tilde{p} \rangle \quad \text{for all } \tilde{p} \in \mathcal{V}.$$

Similarly, the action of the Hessian,  $\Theta_{mm}(\bar{m})$  in a direction  $\zeta \in \mathcal{H}$ , can be expressed as

$$(30) \quad \Theta_{mm}(\bar{m})\zeta = e^{\bar{m}}(\zeta \nabla u \cdot \nabla p + \nabla \cdot \zeta \nabla p + \nabla u \cdot \nabla \zeta),$$

where  $u$  and  $p$  are the state and adjoint variables computed with a given control  $z$  at  $m = \bar{m}$ . The variables  $\zeta$  and  $\tilde{\zeta}$ , which we refer to as the *incremental* state and adjoint variables, are obtained by solving the following incremental state and adjoint equations:

$$(31a) \quad \langle e^{\bar{m}} \nabla \zeta, \nabla \tilde{\zeta} \rangle = -\langle \zeta e^{\bar{m}} \nabla u, \nabla \tilde{\zeta} \rangle \quad \text{for all } \tilde{\zeta} \in \mathcal{V},$$

$$(31b) \quad \langle e^{\bar{m}} \nabla \tilde{\zeta}, \nabla \zeta \rangle = -\langle \mathcal{Q}^* \mathcal{Q} \zeta, \tilde{\zeta} \rangle - \langle \zeta e^{\bar{m}} \nabla p, \nabla \tilde{\zeta} \rangle \quad \text{for all } \tilde{\zeta} \in \mathcal{V}.$$

Note that (28) and (30) for  $\Theta_m$  and  $\Theta_{mm}$  require that the state and adjoint equations (25) and (29), as well as their incremental variants (31), are satisfied. Thus, in the formulation of the OUU objective function (18a), these equations must be enforced as constraints.

**5.1. The OUU problem for (25).** Here we summarize the formulation of the OUU problem, which uses a quadratic approximation of  $\Theta$  defined in (27):

$$(32a) \quad \min_{z \in \mathcal{Z}} \mathcal{J}(z) := \frac{1}{2} \|\mathcal{Q}u - \bar{\mathbf{q}}\|_2^2 + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, \psi_j \rangle + \frac{1}{2} \left\{ \langle \Theta_m(\bar{m}), \mathcal{C}[\Theta_m(\bar{m})] \rangle + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \psi_j, \mathcal{C}\psi_j \rangle \right\} + \frac{\gamma}{2} \|z\|^2,$$

where for  $j = 1, \dots, n_{\text{tr}}$ ,

$$(32b) \quad \Theta_m(\bar{m}) = e^{\bar{m}} \nabla u \cdot \nabla p,$$

$$(32c) \quad \psi_j = e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \zeta_j \nabla p + \nabla u \cdot \nabla \zeta_j),$$

and the variables  $(u, p, \{\zeta_j\}, \{\tilde{\zeta}_j\}) \in \mathcal{V}^2 \times (\mathcal{V}^{n_{\text{tr}}})^2$ , which can be considered the state variables of the OUU problem, solve

$$(32d) \quad \langle e^{\bar{m}} \nabla u, \nabla \tilde{u} \rangle = \langle b + Fz, \tilde{u} \rangle \quad \text{for all } \tilde{u} \in \mathcal{V},$$

$$(32e) \quad \langle e^{\bar{m}} \nabla p, \nabla \tilde{p} \rangle = -\langle \mathcal{Q}^*(\mathcal{Q}u - \bar{\mathbf{q}}), \tilde{p} \rangle \quad \text{for all } \tilde{p} \in \mathcal{V},$$

$$(32f) \quad \langle e^{\bar{m}} \nabla \zeta_j, \nabla \tilde{\zeta} \rangle = -\langle \zeta_j e^{\bar{m}} \nabla u, \nabla \tilde{\zeta} \rangle \quad \text{for all } \tilde{\zeta} \in \mathcal{V},$$

$$(32g) \quad \langle e^{\bar{m}} \nabla \tilde{\zeta}_j, \nabla \zeta \rangle = -\langle \mathcal{Q}^* \mathcal{Q} \zeta_j, \tilde{\zeta} \rangle - \langle \zeta_j e^{\bar{m}} \nabla p, \nabla \tilde{\zeta} \rangle \quad \text{for all } \tilde{\zeta} \in \mathcal{V}.$$

**5.2. Evaluation and gradient computation of the OUU objective function.** Solving the PDE-constrained optimization problem (32) efficiently requires gradient-based optimization methods. To compute the gradient of  $\mathcal{J}$  with respect to  $z$ , we follow a Lagrangian approach and employ adjoint variables (i.e., Lagrange multipliers) to enforce the PDE constraints (32d)–(32g). The details of the derivation of the gradient are relegated to Appendix B. The expression for the gradient, in a direction  $\tilde{z}$ , takes the form

$$(33) \quad \mathcal{G}(z)\tilde{z} = \gamma \langle z, \tilde{z} \rangle - \langle F\tilde{z}, u^* \rangle, \quad \tilde{z} \in L^2(\mathcal{D}),$$

where  $u^*$  is obtained by solving the following system of equations for the OUU adjoint variables  $(u^*, p^*, \{\tilde{u}_j^*, \tilde{p}_j^*\}) \in \mathcal{V}^2 \times (\mathcal{V}^{n_{\text{tr}}})^2$ :

$$(34a) \quad \langle e^{\bar{m}} \nabla \tilde{u}_j^*, \nabla \tilde{u} \rangle = \langle b_1^{(j)}, \tilde{u} \rangle,$$

$$(34b) \quad \langle e^{\bar{m}} \nabla \tilde{p}_j^*, \nabla \tilde{p} \rangle + \langle \mathcal{Q}^* \mathcal{Q} \tilde{u}_j^*, \tilde{u} \rangle = \langle b_2^{(j)}, \tilde{u} \rangle,$$

$$(34c) \quad \langle e^{\bar{m}} \nabla p^*, \nabla \tilde{p} \rangle + \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j e^{\bar{m}} \nabla \tilde{u}_j^*, \nabla \tilde{p} \rangle = \langle b_3, \tilde{p} \rangle,$$

$$(34d) \quad \langle e^{\bar{m}} \nabla u^*, \nabla \tilde{u} \rangle + \langle \mathcal{Q}^* \mathcal{Q} p^*, \tilde{u} \rangle + \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j e^{\bar{m}} \nabla \tilde{u}_j^*, \nabla \tilde{u} \rangle = \langle b_4, \tilde{u} \rangle$$

for all  $(\tilde{u}, \tilde{p}, \tilde{u}, \tilde{p}) \in \mathcal{V}^4$ , with  $\{b_1^{(j)}\}_{j=1}^{n_{\text{tr}}}$ ,  $\{b_2^{(j)}\}_{j=1}^{n_{\text{tr}}}$ ,  $b_3$ , and  $b_4$  given by

$$\begin{aligned} \langle b_1^{(j)}, \tilde{u} \rangle &= - \left[ \frac{1}{2n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}} \nabla u \cdot \nabla \tilde{u} \rangle + \frac{1}{2n_{\text{tr}}} \langle e^{\bar{m}} \nabla u \cdot \nabla \tilde{u}, \mathcal{C}\psi_j \rangle \right], \\ \langle b_2^{(j)}, \tilde{u} \rangle &= - \left[ \frac{1}{2n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}} \nabla \tilde{u} \cdot \nabla p \rangle + \frac{1}{2n_{\text{tr}}} \langle e^{\bar{m}} \nabla \tilde{u} \cdot \nabla p, \mathcal{C}\psi_j \rangle \right], \\ \langle b_3, \tilde{p} \rangle &= - \left[ \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}} (\zeta_j \nabla u \cdot \nabla \tilde{p} + \nabla \tilde{u}_j \cdot \nabla \tilde{p}) \rangle + \langle e^{\bar{m}} \nabla u \cdot \nabla \tilde{p}, \mathcal{C}[\Theta_m(\bar{m})] \rangle \right. \\ &\quad \left. + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle e^{\bar{m}} (\zeta_j \nabla u \cdot \nabla \tilde{p} + \nabla \tilde{u}_j \cdot \nabla \tilde{p}), \mathcal{C}\psi_j \rangle \right], \\ \langle b_4, \tilde{u} \rangle &= - \left[ \langle \mathcal{Q}^* (\mathcal{Q}u - \bar{\mathbf{q}}), \tilde{u} \rangle + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}} (\zeta_j \nabla \tilde{u} \cdot \nabla p + \nabla \tilde{u} \cdot \nabla \tilde{u}_j) \rangle \right. \\ &\quad \left. + \langle e^{\bar{m}} \nabla \tilde{u} \cdot \nabla p, \mathcal{C}[\Theta_m(\bar{m})] \rangle \right. \\ &\quad \left. + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle e^{\bar{m}} (\zeta_j \nabla \tilde{u} \cdot \nabla p + \nabla \tilde{u} \cdot \nabla \tilde{u}_j), \mathcal{C}\psi_j \rangle \right]. \end{aligned}$$

Next, we count the number of PDE solves required for the evaluation of the objective function (32a) and of its gradient (33). Although in the present example, the application of  $\mathcal{C}$  requires PDE solves that are similar to those in the state and adjoint equations, we do not include

them in our counting since (1) the matrix in  $\mathcal{C}$  does not change, and thus a sparse factorization or multigrid solver can be set up up front, and (2) we aim at problems where the state (and thus the adjoint) equation is more complex than in the present case, and thus its solution dominates the application of  $\mathcal{C}$ .

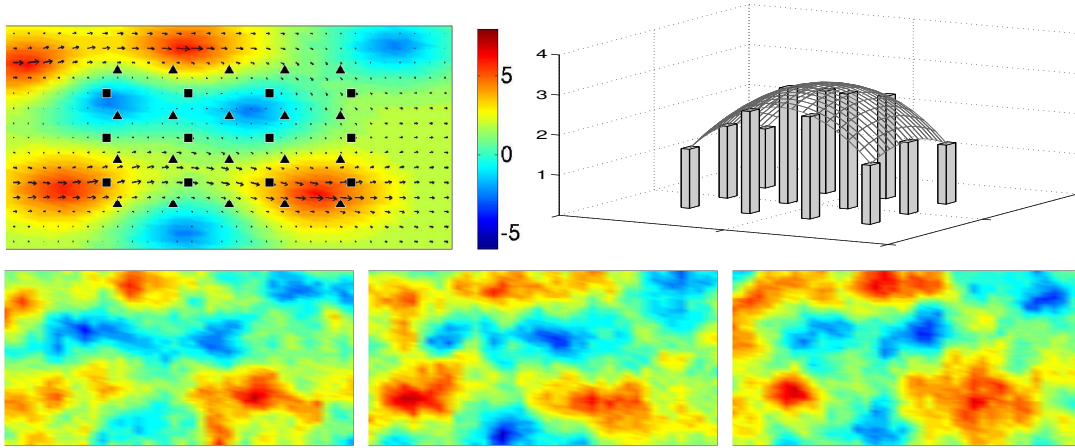
Hence, the number of forward-like PDE solves (i.e., the state equations, the adjoint equation, and incremental equations) required for evaluating the objective function (32a) is  $2 + 2 \times n_{\text{tr}}$ . The gradient evaluation requires another  $2 + 2 \times n_{\text{tr}}$  PDE solves. Thus,  $4 + 4 \times n_{\text{tr}}$  PDE solves are necessary for evaluating the OUU objective function and its gradient. If a small  $n_{\text{tr}}$  provides an approximation of the trace that is suitable for computing an optimal control, each iteration of a gradient-based method requires a moderate, fixed number of PDE solves. We demonstrate this in our numerical experiments, where for a typical OUU problem with a discretized parameter dimension of about 3,000,  $O(10)$  PDE solves are required for approximating the OUU objective and its gradient. This modest computational cost should be contrasted with methods that approximate the OUU objective using sampling or quadrature in the parameter space. While these approaches can, asymptotically, solve the original OUU problem rather than the formulation based on the quadratic expansion of  $\Theta$ , their cost in terms of PDE solves can be much higher. Moreover, if a factorization-based direct solver can be used for the governing PDE problems, these factorizations can be reused multiple times in the quadratic approximation approach. This is the case since the PDE operators arising in each quadratic approximation correspond to the same parameter. Thus, reusing factorizations can save significant computation time. Since each sample in the Monte Carlo approach corresponds to a different parameter, matrix factorizations must be recomputed for each sample.

The computational cost of sampling-based methods is exacerbated for OUU problems governed by nonlinear forward PDEs; in such cases, while our approach requires only one nonlinear PDE solve and  $\mathcal{O}(n_{\text{tr}})$  linear(ized) PDE solves (for OUU objective and gradient evaluation), the required number of nonlinear PDE solves for sampling approaches scales with the number of samples.

**6. Computational experiments.** Next, we numerically study our OUU approach applied to the control of an elliptic PDE with an uncertain coefficient, as discussed in section 5.

**6.1. Problem setup.** We consider a rectangular domain  $\mathcal{D} = (0, 2) \times (0, 1) \subset \mathbb{R}^2$  for (25), impose Dirichlet boundary conditions on the left and right sides of the domain according to  $u(0, y) = 1$  and  $u(2, y) = 0$  for  $y \in [0, 1]$ , and use homogeneous Neumann boundary conditions on the top and bottom boundaries. We use a finite element mesh with 3,081 linear rectangular elements and 3,200 degrees of freedom to discretize the state and adjoint variables and the uncertain parameter field. The discretized uncertain parameters are the coefficients in the finite element expansion of the parameter field.

*The probability law of the uncertain parameter field.* The probability law of the uncertain parameter, which is the log-permeability field  $m$ , is given by a Gaussian measure  $\mu = \mathcal{N}(\bar{m}, \mathcal{C})$ . The mean  $\bar{m}$  and the locations of production and control wells are shown in Figure 1 (upper left image). The covariance operator  $\mathcal{C}$  of  $\mu$  is as in (8), with  $\kappa = 2 \times 10^{-2}$  and  $\alpha = 4$ . These values are chosen such that samples from  $\mu$  have a desired magnitude and correlation length. Typical realizations of the random process  $m$  are also shown in Figure 1.



**Figure 1.** Top row: Shown on the left is the mean log-permeability field with corresponding Darcy flow driven by the lateral boundary conditions. Superimposed is the well configuration (triangles indicate the locations of injection/control wells, and squares of production wells). Shown on the right is the target pressure at the production wells, located at points  $\{\mathbf{x}_i\}_{i=1}^{12}$ , computed according to  $\bar{q}_i = q(\mathbf{x}_i)$ , where  $q(\mathbf{x}) = 3 - 4(x_1 - 1)^2 - 8(x_2 - 0.5)^2$ . Bottom row: Realizations of the uncertain log-permeability field.

**The control.** For the right-hand side in (25), we consider  $b \equiv 0$  and define  $F$  as a weighted sum of finitely many mollified Dirac delta functions, which we denote by  $f_i$ ,  $i = 1, \dots, n_c$ :

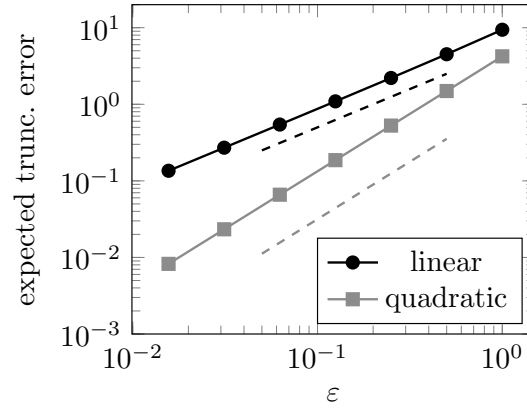
$$(35) \quad F\mathbf{z} = \sum_{i=1}^{n_c} z_i f_i.$$

The weights  $z_i$  in (35) are the (finitely many) control variables, summarized in the control vector  $\mathbf{z} = (z_1, \dots, z_{n_c})^T$ . The right-hand side (35) is motivated by subsurface flow applications in petroleum engineering, where fluid injection at injection wells is used to control the flow rates at production wells. Here,  $f_i = f_i(x, y)$  represent the locations of injection wells, and  $z_i$  are the injection rates at these wells. We impose control bounds; i.e., the set of admissible controls is

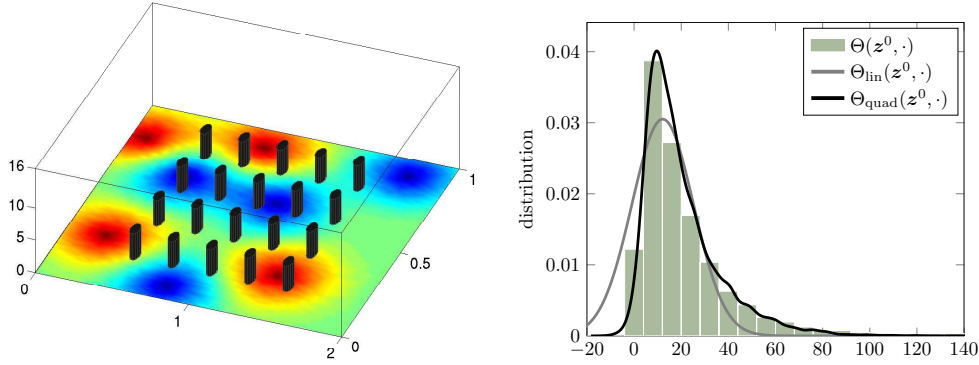
$$\mathcal{L}_{\text{ad}} = \{\mathbf{z} \in \mathbb{R}^{n_c} : z_{\min} \leq z_i \leq z_{\max}, 1 \leq i \leq n_c\},$$

where  $z_{\min} = 0$  and  $z_{\max} = 16$ . In the present example, we use  $n_c = 20$  control wells; see Figure 1 (top left). The control objective (27) is the squared  $\ell^2$  difference between the pressure at the production wells and a vector  $\bar{\mathbf{q}}$  of target pressure values, which follows a parabolic profile as depicted in Figure 1 (top right).

**Quadratic approximations in the small variance limit.** Here, we compare approximation error in the linear and quadratic approximations of the parameter-to-objective map, at a fixed control  $\mathbf{z}^0$  with  $z_i^0 = 4$  for  $i = 1, \dots, n_c$ , in the sense discussed in section 3.1.2. In Figure 2, we consider the expected value of the truncation error of the first- and second-order Taylor expansion to the mapping  $m \mapsto \Theta(\mathbf{z}^0, m)$ . This illustrates the rate of decay of the expected truncation error as indicated by the analysis in section 3.1.2. In particular, the truncation error in a linear expansion is  $\mathcal{O}(\varepsilon)$ , and that of the quadratic approximation is  $\mathcal{O}(\varepsilon^{3/2})$ . In this study, we consider scaling the distribution law of  $m$  given according to  $\mu = \mathcal{N}(\bar{m}, \varepsilon\mathcal{C})$  with successively smaller values of  $\varepsilon$ .



**Figure 2.** The expected values  $E\{|\Theta(z, \cdot) - \Theta_{\text{lin}}(z, \cdot)|\}$  (black solid line) and  $E\{|\Theta(z, \cdot) - \Theta_{\text{quad}}(z, \cdot)|\}$  (gray solid line), with the distribution law of  $m$  given by  $\mu = \mathcal{N}(\bar{m}, \varepsilon \mathcal{C})$  as  $\varepsilon \rightarrow 0$ . The black and gray dashed lines indicate  $\mathcal{O}(\varepsilon)$  and  $\mathcal{O}(\varepsilon^{3/2})$ , respectively. The errors are computed for  $\varepsilon_k = 1/2^k$ , for  $k = 0, \dots, 6$ , using a fixed Monte Carlo sample of size 10,000 to approximate the average errors.



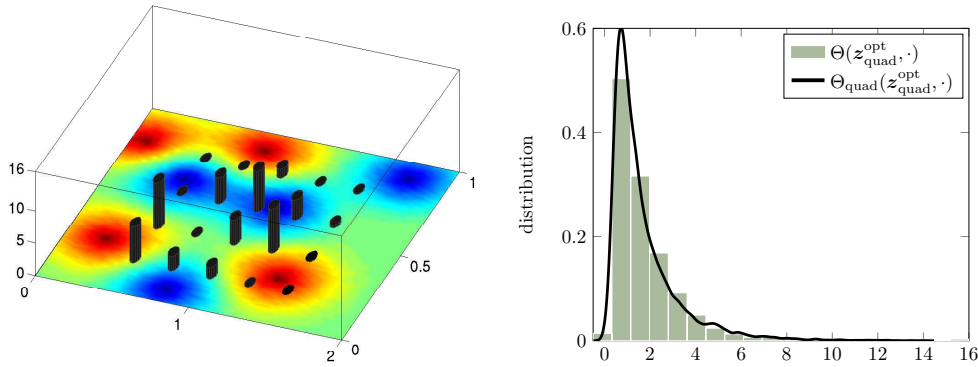
**Figure 3.** Left: Control  $\mathbf{z}^0$  with constant values  $z_i^0 = 4$ ,  $i \in \{1, \dots, n_c\}$  (in black). Right: Histogram of  $\Theta(\mathbf{z}^0, m)$  and the distributions of  $\Theta_{\text{lin}}$  and  $\Theta_{\text{quad}}$ , which are based on first- and second-order expansions of the parameter-to-objective function (27).

**6.2. Results.** We solve the numerical optimization problem (32) via an interior point method (to incorporate the box constraints for  $\mathbf{z}$ ) with BFGS Hessian approximation. For that purpose, we employ the MATLAB `fmincon` routine, to which we provide functions for the evaluation of the objective function (32) and for the computation of its gradient with respect to the control, as derived in the previous section.

We use a uniform control vector  $\mathbf{z}^0 = (4, \dots, 4)$  (see Figure 3, left) as the initial guess of the optimization algorithm. It is instructive to consider the statistical distribution of the control objective (as defined in (27)) for this initialization, which is shown in the right image in Figure 3. We also depict the distributions of the linear and quadratic approximations  $\Theta_{\text{lin}}(\mathbf{z}^0, m)$  and  $\Theta_{\text{quad}}(\mathbf{z}^0, m)$  as defined in (4) and (5), respectively. For all figures, we have used kernel density estimation (KDE) to approximate the probability density functions (PDFs) from 10,000 samples. Note that compared to  $\Theta_{\text{lin}}$ , the distribution of  $\Theta_{\text{quad}}$  is a significantly

better approximation for the distribution of  $\Theta$ .

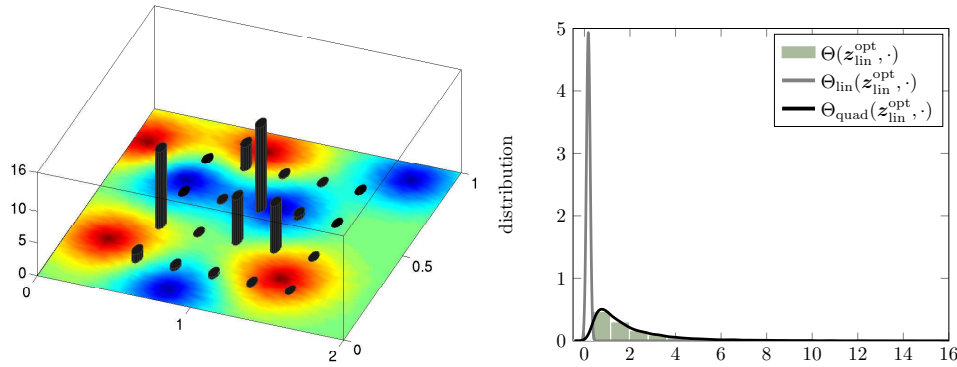
In Figure 4 we show the risk-averse optimal control for the risk-aversion parameter  $\beta = 1$ , the control cost weight  $\gamma = 10^{-5}$ , and where we use  $n_{tr} = 40$  trace estimator vectors for approximating the traces in the OUU objective function. To cope with the nonconvexity of the OUU objective function, we use a continuation strategy with respect to  $\beta$ ; i.e., we solve a sequence of optimization problems with increasing  $\beta$ , namely  $\{\beta_k\}_{k=1}^5 = \{0, 0.25, 0.5, 0.75, 1\}$ . The average number of quasi-Newton interior point iterations required to decrease the residual by  $5 \times 10^{-4}$  was 65. The left figure shows the magnitude of the control vector at each injection well, and the right figure depicts the statistical distribution of the control objective at the computed optimal control for both  $\Theta$  and  $\Theta_{quad}$  (note that the optimal control is based on the latter). To assess how successful the optimal control is in reducing the mean and the standard deviation, compare the right images in Figures 3 and 4. Notice that compared to the constant control  $\mathbf{z}^0$ , the optimal control results in a distribution of the objective that both is shifted to the left (reduction in the mean) and has less spread (reduction in variance).



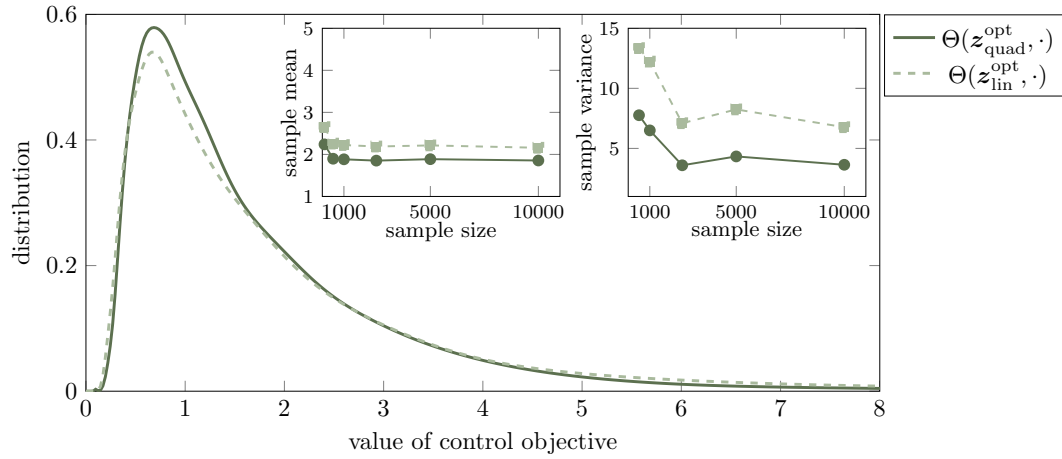
**Figure 4.** Left: The optimal control  $\mathbf{z}_{quad}^{opt}$  computed by minimizing the OUU objective function with quadratic approximation to the parameter-to-objective map (27), where we use  $\beta = 1, \gamma = 10^{-5}$ , and  $n_{tr} = 40$ . The initial guess for the optimizer was  $\mathbf{z}^0$  defined in Figure 3. Right: The distributions of  $\Theta$  and  $\Theta_{quad}$  for the optimal control  $\mathbf{z}_{quad}^{opt}$ .

Next, we study whether using the linear (rather than the quadratic) approximation to the parameter-to-objective map for the computation of risk-averse optimal control can lead to suboptimal results. Using  $\Theta_{lin}$  instead of  $\Theta_{quad}$  results in a simplified version of (32a), obtained by neglecting terms involving the Hessian  $\Theta_{mm}$ . We solve the same risk-averse OUU problem as before but with  $\Theta_{lin}$  rather than  $\Theta_{quad}$ . The resulting optimal control  $\mathbf{z}_{lin}^{opt}$  is shown on the left of Figure 5, and the distributions of  $\Theta(\mathbf{z}_{lin}^{opt}, m)$ ,  $\Theta_{lin}(\mathbf{z}_{lin}^{opt}, m)$ , and  $\Theta_{quad}(\mathbf{z}_{lin}^{opt}, m)$  are shown on the right. Note the large discrepancy between the distributions; in particular, the distribution based on  $\Theta_{lin}$  is a poor approximation to the actual distribution. Our goal, however, is to find a control such that the distribution of the control objective  $\Theta$  has small mean and variance. In the present example, the optimal control computed with the linear approximation of the parameter-to-objective function  $\Theta_{lin}$  does not perform much worse than the optimal control found with  $\Theta_{quad}$  in terms of reducing the mean and variance of the distribution of  $\Theta$ —despite the poor approximation shown in Figure 5. This can be seen from





**Figure 5.** Left: Optimal control  $z_{\text{lin}}^{\text{opt}}$  computed by minimizing the OUU objective function with linear approximation to the parameter-to-objective map, where we use  $\beta = 1, \gamma = 10^{-5}$ . Right: Distributions of the control objective and its approximations for  $z = z_{\text{lin}}^{\text{opt}}$ .

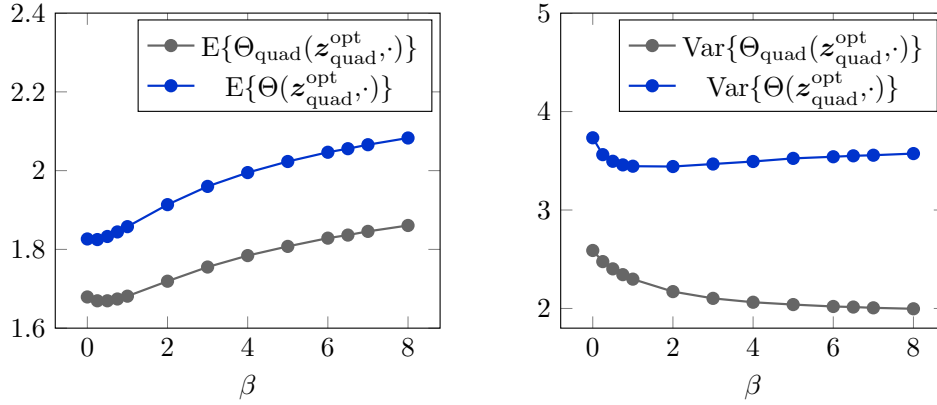


**Figure 6.** Comparison of the distributions of  $\Theta(z, \cdot)$  for  $z = z_{\text{lin}}^{\text{opt}}$  (dashed) and  $z = z_{\text{quad}}^{\text{opt}}$  (solid) for  $\beta = 1, \gamma = 10^{-5}$ , and  $n_{\text{tr}} = 40$  trace estimation vectors. The inserts show the Monte Carlo sample convergence for the mean and the variance of the distributions.

Figure 6, where we compare the distribution of  $\Theta$  for the optimal controls  $z_{\text{lin}}^{\text{opt}}$  and  $z_{\text{quad}}^{\text{opt}}$ . We notice that the distribution of  $\Theta(z_{\text{lin}}^{\text{opt}}, m)$  has slightly larger mean and variance, as can be observed by the thicker tail of the distribution.

**Influence of risk-aversion parameter on the optimal control.** Next, we study the effect of the parameter  $\beta$  in (32a) on the optimal control and the corresponding distribution of the control objective. In Figure 7, we show results for the mean and the variance for various risk-aversion parameters  $\beta$ . While the optimal controls have been computed using  $\Theta_{\text{quad}}$ , we also report the mean and variance of  $\Theta$ . With increasing  $\beta$ , the mean of  $\Theta_{\text{quad}}$  increases and the variance of  $\Theta_{\text{quad}}$  decreases, as expected. The optimal controls were computed using a fixed randomized trace estimator with  $n_{\text{tr}} = 40$ .

Next, we consider the effect of increasing  $\beta$  on the distribution of  $\Theta$  itself. While for small

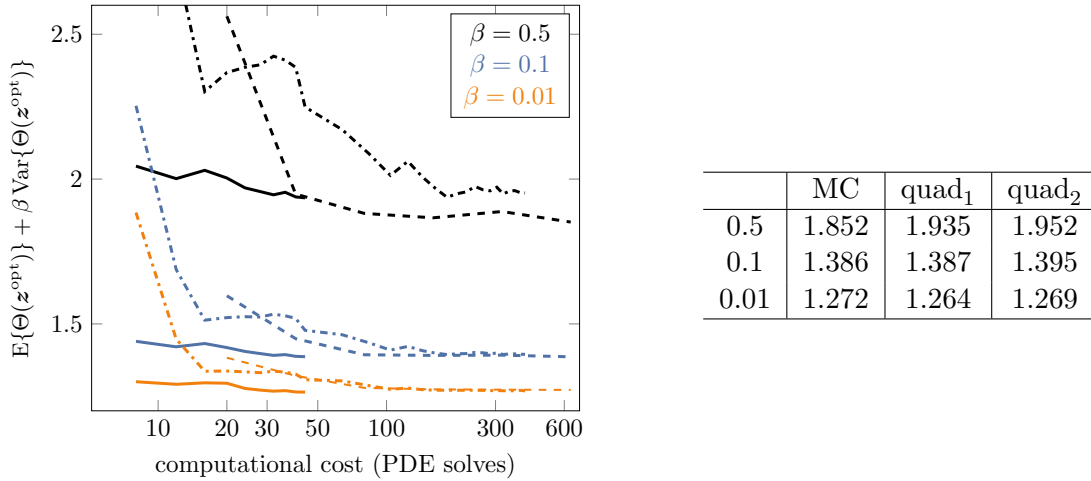


**Figure 7.** Effect of risk-aversion parameter  $\beta$  on the mean and the variance of the control objective (27) and its quadratic approximation at the optimal control  $z_{\text{quad}}^{\text{opt}}$ . We report statistical values for  $\Theta_{\text{quad}}$ , which has been used in the computation of the optimal control, as well as statistical values for the true objective  $\Theta$ . These plots were generated using Monte Carlo sampling, with 10,000 sample points.

values, increasing  $\beta$  results in smaller variances for  $\Theta(z, m)$ , as desired, for larger values of  $\beta$ , the variance increases with  $\beta$ . This can be attributed to the fact that for computing the optimal control, we use the quadratic approximation  $\Theta_{\text{quad}}(z, m)$ , which only approximates the moments of the control objective  $\Theta(z, m)$ .

**Comparison between quadratic approximation and Monte Carlo.** Next, we compare the computational cost for computing the optimal controls using the approach based on the quadratic approximation versus an SAA approach, where Monte Carlo sampling is used to approximate (2). In Figure 8 (left), we plot the true OUU objective, approximated accurately using a large number of samples, against the cost per iteration measured in the number of PDE solves required for objective and gradient computation, for three different risk-aversion parameters  $\beta$ . For the quadratic approximation, we approximate the trace terms in the OUU objective function using randomized trace estimation (dash-dotted lines) and using eigenvectors of the covariance-preconditioned Hessian operator (solid lines), as described in section 3.2. For the randomized trace estimation, we use  $n_{\text{tr}}$  values in the range of 1–100 and use the same trace estimator vectors for different values of  $\beta$ ; for the eigenvector-based approach, we use  $n_{\text{tr}} \in \{1, \dots, 10\}$ . The eigenvectors are computed for a reference (namely, the initial) control. The Monte Carlo sampling approach is shown with a dashed line; for this approach, we use the same sequence of samples for the different values of  $\beta$ , with samples of size 10, 20, 40, 80, 160, and 320. Note that in the Monte Carlo approach, the cost per iteration for objective and gradient computation is twice the size of the Monte Carlo sample used.

For a more precise comparison, in Figure 8 (right) for each case we list the data values corresponding to the largest number of PDE solves. The figure on the left shows the following important results: (1) the eigenvector-based trace estimation outperforms the randomized trace estimation for all choices of  $\beta$ ; in both cases, as the number of vectors used for the approximation of the trace increases, the function values level off, indicating convergence of the respective optimal controls; (2) the optimal controls computed using Monte Carlo sampling of the objective asymptotically converge to the exact optimal controls of the true OUU objective;



**Figure 8.** Left: OUU objective evaluated at differently computed optimal controls  $\mathbf{z}^{\text{opt}}$  versus the number of PDE solves required for one objective/gradient computation. Shown are results with different risk-aversion parameters  $\beta$  for optimal controls computed using a quadratic approximation (solid and dash-dotted lines) and Monte Carlo sampling (dashed lines). For the Monte Carlo sampling, the same sequence of samples was used for the different values of  $\beta$ . For the quadratic approximation, to estimate the trace we used both the basis of eigenvectors of the covariance-preconditioned Hessian, evaluated at the initial point  $\mathbf{z}_0$  (solid line) and random vectors (dash-dotted line). Right: The OUU objective values corresponding to the largest number of PDE solves for various  $\beta$  values. The columns labeled with MC, quad<sub>1</sub>, and quad<sub>2</sub> correspond to the Monte Carlo sampling, quadratic approximation with eigenvector- and random vectors-based trace estimations, respectively. To approximate the OUU objective at the computed optimal controls, we use a fixed Monte Carlo sample of size 10,000.

hence, these controls will asymptotically outperform optimal controls based on the quadratic approximation; and (3) for low risk-aversion (i.e., small  $\beta$ ) or for limited compute time (i.e., when only a small number of PDE solves is afforded), the quadratic approximation approach is superior to the Monte Carlo SAA approach.

**7. Conclusions.** We propose a scalable method for risk-averse optimal control of systems governed by PDEs with uncertain parameter fields. Our approach uses a quadratic approximation of the parameter-to-objective map, which enables one to analytically compute the moments appearing in the OUU objective function. Moreover, we employ randomized trace estimators for the operator traces in the OUU objective function. The resulting optimization problem is constrained by the PDEs defining the gradient  $\Theta_m$  and the linear action of the Hessian  $\Theta_{mm}$ . The resulting method for risk-averse OUU is applicable to problems with high-dimensional discretized parameter spaces. This is demonstrated in numerical tests, where we present results for a problem with a 3,200-dimensional (discretized) parameter space. Hence, our approach provides a practical alternative to computationally expensive sampling-based OUU methods. The advantages of our method compared to sampling/quadrature methods are even more pronounced in the context of risk-averse optimal control of systems governed by *nonlinear* PDEs. Whereas our approach requires only one nonlinear PDE solve and about  $2 \times n_{\text{tr}}$  linear(ized) PDE solves (to evaluate the OUU objective function), the required number

of nonlinear PDE solves for sampling approaches scales with the number of samples.

**Appendix A. Expectation of the truncation error in a Taylor expansion.** We begin by stating a well-known result [43, Theorem 6.6] regarding Gaussian measures on a Hilbert space.

**Theorem A.1.** *Let  $\mu = \mathcal{N}(0, \mathcal{C})$  be a Gaussian measure on a Hilbert space  $\mathcal{H}$ . For any integer  $r \geq 1$ , there is a constant  $B = B_r \geq 0$  such that*

$$(36) \quad \int_{\mathcal{H}} \|x\|^{2r} \mu(dx) \leq B \text{Tr}(\mathcal{C})^r.$$

Note that the above theorem, for the case of  $r = 2$ , follows from (15). Next, we define the notation  $T[m^p] := T[m, m, \dots, m]$  for the action of a  $p$ -linear map  $T : \mathcal{H} \times \mathcal{H} \times \dots \times \mathcal{H} \rightarrow \mathbb{R}$ .

**Lemma A.2.** *Let  $\mu = \mathcal{N}(0, \mathcal{C})$ , and suppose  $T$  is a symmetric  $p$ -linear map, with  $p \geq 2$ , such that  $|T[m^p]| \leq K \|m\|^p$  for all  $m \in \mathcal{H}$ . Then we have  $\int_{\mathcal{H}} |T[m^p]| \mu(dm) \leq \tilde{K} \text{Tr}(\mathcal{C})^{p/2}$  for a positive constant  $\tilde{K}$ .*

*Proof.* For  $p = 2$ , the proof is straightforward. For  $p > 2$ , if  $p$  is even, i.e.,  $p = 2k$  for  $k \in \mathbb{Z}_+$ , then  $\int_{\mathcal{H}} |T[m^{2k}]| \mu(dm) \leq K \int_{\mathcal{H}} \|m\|^{2k} \mu(dm) \leq \tilde{K} \text{Tr}(\mathcal{C})^k$ , where the last inequality follows from Theorem A.1. Note that here  $\tilde{K} = KB$  with  $B$  from (36). In the case  $p = 2k + 1$  for  $k \in \mathbb{Z}_+$ , we have

$$\begin{aligned} \int_{\mathcal{H}} |T[m^{2k+1}]| \mu(dm) &\leq K \int_{\mathcal{H}} \|m\|^{2k+1} \mu(dm) \\ &\leq K \left[ \int_{\mathcal{H}} \|m\|^2 \mu(dm) \right]^{1/2} \left[ \int_{\mathcal{H}} \|m\|^{4k} \mu(dm) \right]^{1/2} \leq \tilde{K} \text{Tr}(\mathcal{C})^{1/2} \text{Tr}(\mathcal{C})^k = \tilde{K} \text{Tr}(\mathcal{C})^{(2k+1)/2}. \blacksquare \end{aligned}$$

**Proposition A.3.** *Let  $\Theta : (\mathcal{H}, \mathcal{B}(\mathcal{H}), \mu) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  be, almost surely, a  $p + 1$  times continuously differentiable function, and assume  $m = \mathcal{N}(\bar{m}, \varepsilon \mathcal{C})$ , with  $\varepsilon > 0$ . Suppose  $\Theta$  has  $p + 1$  uniformly bounded derivatives. Then the expected value of the truncation error of the  $p$ th-order Taylor expansion is  $\mathcal{O}(\varepsilon^{(p+1)/2})$ .*

*Proof.* Since  $\Theta : \mathcal{H} \rightarrow \mathbb{R}$  is  $p + 1$  times continuously differentiable, we have

$$\Theta(m) = \Theta(\bar{m}) + \sum_{n=1}^p \frac{1}{n!} \Theta^{(n)}(\bar{m})(m - \bar{m})^n + R_p(m; \bar{m}).$$

The remainder term is given by  $R_p(m; \bar{m}) = \frac{1}{(p+1)!} \Theta^{(p+1)}(\xi)[(m - \bar{m})^{p+1}]$ , where  $\xi$  is in the interior of the line segment between  $m$  and  $\bar{m}$ . Here we use the notation  $\Theta^{(n)}$  for the  $n$ th derivative. Now by assumption of the proposition,  $\Theta^{(k+1)}$  is uniformly bounded on  $\mathcal{H}$ . Therefore, by Lemma A.2 we get  $\mathbb{E}\{R_p(m; \bar{m})\}$  is  $\mathcal{O}(\varepsilon^{(p+1)/2})$ .  $\blacksquare$

**Appendix B. Derivation of the gradient of the OUU objective function  $\mathcal{J}$ .** Here, we summarize the derivation of the gradient of the OUU objective function presented in (32a). To derive the expression for the gradient, we employ a formal Lagrangian approach [45, 8], which uses a Lagrangian function composed of the objective function (32a) with the PDE

constraints (32d)–(32g) enforced through Lagrange multiplier functions. This Lagrangian function  $\mathcal{L}$  for the OUU problem is given by

$$\begin{aligned} \mathcal{L}(z, u, p, \{ \cdot \}_j, \{ \cdot \}_j, u^*, p^*, \{ \cdot \}_j^*, \{ \cdot \}_j^*) \\ = \frac{1}{2} \|\mathcal{Q}u - \bar{\mathbf{q}}\|_2^2 + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \nabla_j p + \nabla u \cdot \nabla \cdot \nabla_j) \rangle \\ + \frac{1}{2} \langle e^{\bar{m}} \nabla u \cdot \nabla p, \mathcal{C}[e^{\bar{m}} \nabla u \cdot \nabla p] \rangle + \frac{1}{4n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \left\| \mathcal{C}^{1/2}[e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \nabla_j p + \nabla u \cdot \nabla \cdot \nabla_j)] \right\|^2 \\ + \frac{\gamma}{2} \|z\|^2 + \langle e^{\bar{m}} \nabla u, \nabla u^* \rangle - \langle b + Fz, u^* \rangle + \langle e^{\bar{m}} \nabla p, \nabla p^* \rangle + \langle \mathcal{Q}^*(\mathcal{Q}u - \bar{\mathbf{q}}), p^* \rangle \\ + \sum_{j=1}^{n_{\text{tr}}} \left[ \langle e^{\bar{m}} \nabla \cdot \nabla_j, \nabla \cdot \nabla_j^* \rangle + \langle \zeta_j e^{\bar{m}} \nabla u, \nabla \cdot \nabla_j^* \rangle \right] \\ + \sum_{j=1}^{n_{\text{tr}}} \left[ \langle e^{\bar{m}} \nabla \cdot \nabla_j, \nabla \cdot \nabla_j^* \rangle + \langle \mathcal{Q}^* \mathcal{Q} \cdot \nabla_j, \nabla \cdot \nabla_j^* \rangle + \langle \zeta_j e^{\bar{m}} \nabla p, \nabla \cdot \nabla_j^* \rangle \right]. \end{aligned}$$

The variables  $(u, p, \{ \cdot \}_j, \{ \cdot \}_j) \in \mathcal{V}^2 \times (\mathcal{V}^{n_{\text{tr}}})^2$  are the OUU state variables, and  $(u^*, p^*, \{ \cdot \}_j^*, \{ \cdot \}_j^*) \in \mathcal{V}^2 \times (\mathcal{V}^{n_{\text{tr}}})^2$  are the OUU adjoint variables, with  $j \in \{1, \dots, n_{\text{tr}}\}$ . Requiring that variations of  $\mathcal{L}$  with respect to the OUU adjoint variables vanish, we recover the OUU state equations (32d)–(32g). The variations of  $\mathcal{L}$  with respect to the OUU state variables are

$$\begin{aligned} \mathcal{L}_u[\tilde{u}] &= \langle \mathcal{Q}^*(\mathcal{Q}u - \bar{\mathbf{q}}), \tilde{u} \rangle + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}}(\zeta_j \nabla \tilde{u} \cdot \nabla p + \nabla \tilde{u} \cdot \nabla \cdot \nabla_j) \rangle + \langle e^{\bar{m}} \nabla \tilde{u} \cdot \nabla p, \mathcal{C}[\Theta_m(\bar{m})] \rangle \\ &\quad + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle e^{\bar{m}}(\zeta_j \nabla \tilde{u} \cdot \nabla p + \nabla \tilde{u} \cdot \nabla \cdot \nabla_j), \mathcal{C}[e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \nabla_j p + \nabla u \cdot \nabla \cdot \nabla_j)] \rangle \\ &\quad + \langle e^{\bar{m}} \nabla \tilde{u}, \nabla u^* \rangle + \langle \mathcal{Q}^* \mathcal{Q} \tilde{u}, p^* \rangle + \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j e^{\bar{m}} \nabla \tilde{u}, \nabla \cdot \nabla_j^* \rangle, \\ \mathcal{L}_p[\tilde{p}] &= \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla \tilde{p} + \nabla \cdot \nabla_j \tilde{p}) \rangle + \langle e^{\bar{m}} \nabla u \cdot \nabla \tilde{p}, \mathcal{C}[\Theta_m(\bar{m})] \rangle \\ &\quad + \frac{1}{2n_{\text{tr}}} \sum_{j=1}^{n_{\text{tr}}} \langle e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla \tilde{p} + \nabla \cdot \nabla_j \tilde{p}), \mathcal{C}[e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \nabla_j p + \nabla u \cdot \nabla \cdot \nabla_j)] \rangle \\ &\quad + \langle e^{\bar{m}} \nabla \tilde{p}, \nabla p^* \rangle + \sum_{j=1}^{n_{\text{tr}}} \langle \zeta_j e^{\bar{m}} \nabla \tilde{p}, \nabla \cdot \nabla_j^* \rangle, \\ \mathcal{L}_j[\tilde{\cdot}] &= \frac{1}{2n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}} \nabla \tilde{\cdot} \cdot \nabla p \rangle + \frac{1}{2n_{\text{tr}}} \langle e^{\bar{m}} \nabla \tilde{\cdot} \cdot \nabla p, \mathcal{C}[e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \nabla_j p + \nabla u \cdot \nabla \cdot \nabla_j)] \rangle \\ &\quad + \langle e^{\bar{m}} \nabla \tilde{\cdot}, \nabla \cdot \nabla_j^* \rangle + \langle \mathcal{Q}^* \mathcal{Q} \tilde{\cdot}, \nabla \cdot \nabla_j^* \rangle, \\ \mathcal{L}_{\cdot j}[\tilde{\cdot}] &= \frac{1}{2n_{\text{tr}}} \langle \zeta_j, e^{\bar{m}} \nabla u \cdot \nabla \tilde{\cdot} \rangle + \frac{1}{2n_{\text{tr}}} \langle e^{\bar{m}} \nabla u \cdot \nabla \tilde{\cdot}, \mathcal{C}[e^{\bar{m}}(\zeta_j \nabla u \cdot \nabla p + \nabla \cdot \nabla_j p + \nabla u \cdot \nabla \cdot \nabla_j)] \rangle \\ &\quad + \langle e^{\bar{m}} \nabla \tilde{\cdot}, \nabla \cdot \nabla_j^* \rangle, \end{aligned}$$

with  $(\tilde{u}, \tilde{p}, \tilde{\cdot}, \tilde{\cdot}) \in \mathcal{V}^4$ . Letting these variations vanish results in the OUU adjoint equations (34a)–(34d). Finally, the gradient for (32a) is given by

$$(37) \quad \mathcal{L}_z[\tilde{z}] := \mathcal{G}(z)\tilde{z} = \gamma\langle z, \tilde{z} \rangle - \langle F\tilde{z}, u^* \rangle, \quad \tilde{z} \in L^2(\mathcal{D}).$$

## REFERENCES

- [1] A. ALEXANDERIAN, P. J. GLOOR, AND O. GHATTAS, *On Bayesian A- and D-optimal experimental designs in infinite dimensions*, Bayesian Anal., 11 (2016), pp. 671–695.
- [2] A. ALEXANDERIAN, N. PETRA, G. STADLER, AND O. GHATTAS, *A fast and scalable method for A-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems*, SIAM J. Sci. Comput., 38 (2016), pp. A243–A272, <https://doi.org/10.1137/140992564>.
- [3] H. AVRON AND S. TOLEDO, *Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix*, J. ACM (JACM), 58 (2011), 8.
- [4] H.-G. BEYER AND B. SENDHOFF, *Robust optimization—a comprehensive survey*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 3190–3218.
- [5] L. T. BIEGLER, O. GHATTAS, M. HEINKENSCHLOSS, D. KEYES, AND B. VAN BLOEMEN WAANDERS, EDS., *Real-Time PDE-Constrained Optimization*, Comput. Sci. Engrg. 3, SIAM, 2007, <https://doi.org/10.1137/1.9780898718935>.
- [6] J. R. BIRGE AND F. LOUVEAUX, *Introduction to Stochastic Programming*, Springer-Verlag, 1997.
- [7] F. BONIZZONI AND F. NOBILE, *Perturbation analysis for the Darcy problem with log-normal permeability*, SIAM/ASA J. Uncertainty Quant., 2 (2014), pp. 223–244, <https://doi.org/10.1137/130949415>.
- [8] A. BORZÌ AND V. SCHULZ, *Computational Optimization of Systems Governed by Partial Differential Equations*, Comput. Sci. Engrg. 8, SIAM, 2012, <https://doi.org/10.1137/1.9781611972054>.
- [9] A. BORZÌ, V. SCHULZ, C. SCHILLINGS, AND G. VON WINCKEL, *On the treatment of distributed uncertainties in PDE-constrained optimization*, GAMM-Mitt., 33 (2010), pp. 230–246.
- [10] A. BORZÌ AND G. VON WINCKEL, *Multigrid methods and sparse-grid collocation techniques for parabolic optimal control problems with random coefficients*, SIAM J. Sci. Comput., 31 (2009), pp. 2172–2192, <https://doi.org/10.1137/070711311>.
- [11] A. BORZÌ AND G. VON WINCKEL, *A POD framework to determine robust controls in PDE optimization*, Comput. Vis. Sci., 14 (2011), pp. 91–103.
- [12] T. BUI-THANH, O. GHATTAS, J. MARTIN, AND G. STADLER, *A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion*, SIAM J. Sci. Comput., 35 (2013), pp. A2494–A2523, <https://doi.org/10.1137/12089586X>.
- [13] P. CHEN AND A. QUARTERONI, *Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint*, SIAM/ASA J. Uncertainty Quant., 2 (2014), pp. 364–396, <https://doi.org/10.1137/130940517>.
- [14] G. DA PRATO, *An Introduction to Infinite-Dimensional Analysis*, Universitext, Springer, 2006.
- [15] G. DA PRATO AND J. ZABCZYK, *Second-Order Partial Differential Equations in Hilbert Spaces*, Cambridge University Press, 2002.
- [16] M. DAMBRINE, C. DAPOGNY, AND H. HARBRECHT, *Shape optimization for quadratic functionals and states with random right-hand sides*, SIAM J. Control Optim., 53 (2015), pp. 3081–3103, <https://doi.org/10.1137/15M1017041>.
- [17] J. DARLINGTON, C. PANTELIDES, B. RUSTEM, AND B. TANYI, *Decreasing the sensitivity of open-loop optimal solutions in decision making under uncertainty*, European J. Oper. Res., 121 (2000), pp. 343–362.
- [18] J. DARLINGTON, C. C. PANTELIDES, B. RUSTEM, AND B. A. TANYI, *An algorithm for constrained nonlinear optimization under uncertainty*, Automatica, 35 (1999), pp. 217–228.
- [19] M. DASHTI AND A. M. STUART, *The Bayesian approach to inverse problems*, in Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer, 2017, pp. 311–428.
- [20] J. C. DE LOS REYES, *Numerical PDE-constrained optimization*, Springer, 2015.
- [21] M. DIEHL, H. G. BOCK, AND E. KOSTINA, *An approximation technique for robust nonlinear optimization*, Mathematical Programming, 107 (2006), pp. 213–230.



- [22] Y. ERMOLIEV, *Stochastic quasigradient methods and their application to system optimization*, Stochastics, 9 (1983), pp. 1–36.
- [23] X. FERNIQUE, *Intégrabilité des vecteurs gaussiens*, C. R. Acad. Sci. Paris Sér. A-B, 270 (1970), pp. A1698–A1699.
- [24] A. GAIVORONSKII, *Nonstationary stochastic programming problems*, Cybernet. Systems Anal., 14 (1978), pp. 575–579.
- [25] M. D. GUNZBURGER, *Perspectives in Flow Control and Optimization*, Adv. Des. Control 5, SIAM, 2003, <https://doi.org/10.1137/1.9780898718720>.
- [26] M. D. GUNZBURGER AND J. MING, *Optimal control of stochastic flow over a backward-facing step using reduced-order modeling*, SIAM J. Sci. Comput., 33 (2011), pp. 2641–2663, <https://doi.org/10.1137/100817279>.
- [27] M. HINZE, R. PINNAU, M. ULBRICH, AND S. ULBRICH, *Optimization with PDE Constraints*, Springer, 2009.
- [28] L. S. HOU, J. LEE, AND H. MANOUZI, *Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs*, J. Math. Anal. Appl., 384 (2011), pp. 87–103.
- [29] D. KOURI, *An Approach for the Adaptive Solution of Optimization Problems Governed by Partial Differential Equations with Uncertain Coefficients*, Ph.D. thesis, Rice University, 2012.
- [30] D. P. KOURI, *A multilevel stochastic collocation algorithm for optimization of PDEs with uncertain coefficients*, SIAM/ASA J. Uncertainty Quant., 2 (2014), pp. 55–81, <https://doi.org/10.1137/130915960>.
- [31] D. P. KOURI, M. HEINKENSCHLOSS, D. RIDZAL, AND B. VAN BLOEMEN WAANDERS, *Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty*, SIAM J. Sci. Comput., 36 (2014), pp. A3011–A3029, <https://doi.org/10.1137/140955665>.
- [32] 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), pp. A1847–A1879, <https://doi.org/10.1137/120892362>.
- [33] D. P. KOURI AND T. M. SUROWIEC, *Risk-averse PDE-constrained optimization using the conditional value-at-risk*, SIAM J. Optim., 26 (2016), pp. 365–396, <https://doi.org/10.1137/140954556>.
- [34] 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), pp. 2442–2471, <https://doi.org/10.1137/110847597>.
- [35] R. RACKWITZ, *Reliability analysis—a review and some perspectives*, Structural Safety, 23 (2001), pp. 365–395.
- [36] H. ROBBINS AND S. MONRO, *A stochastic approximation method*, Ann. Math. Statist., 22 (1951), pp. 400–407.
- [37] R. T. ROCKAFELLAR AND S. URYASEV, *Optimization of conditional value-at-risk*, J. Risk, 2 (2000), pp. 21–42.
- [38] F. ROOSTA-KHORASANI AND U. ASCHER, *Improved bounds on sample size for implicit matrix trace estimators*, Found. Comput. Math., (2014), pp. 1–26.
- [39] A. RUSZCZYŃSKI AND W. SYSKI, *A method of aggregate stochastic subgradients with on-line stepsize rules for convex stochastic programming problems*, in Stochastic Programming 84 Part II, Springer, 1986, pp. 113–131.
- [40] N. V. SAHINIDIS, *Optimization under uncertainty: State-of-the-art and opportunities*, Comput. Chem. Eng., 28 (2004), pp. 971–983.
- [41] A. SHAPIRO, D. DENTCHEVA, AND A. RUSZCZYŃSKI, *Lectures on Stochastic Programming: Modeling and Theory*, MOS-SIAM Ser. Optim. 9, SIAM, 2009, <https://doi.org/10.1137/1.9780898718751>.
- [42] S. SMOLYAK, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Soviet Math. Dokl., 4 (1963), pp. 240–243.
- [43] A. M. STUART, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.
- [44] 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), pp. 2659–2682, <https://doi.org/10.1137/110835438>.
- [45] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, Grad. Stud. Math. 112, American Mathematical Society, 2010.
- [46] C. WITHERS, *The moments of the multivariate normal*, Bull. Australian Math. Soc., 32 (1985), pp. 103–107.