

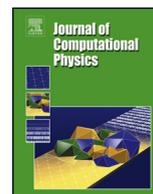


ELSEVIER

Contents lists available at ScienceDirect

## Journal of Computational Physics

www.elsevier.com/locate/jcp



## Taylor approximation and variance reduction for PDE-constrained optimal control under uncertainty

Peng Chen<sup>a,\*</sup>, Umberto Villa<sup>a</sup>, Omar Ghattas<sup>a,b</sup><sup>a</sup> Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX 78712, United States of America<sup>b</sup> Department of Mechanical Engineering, and Department of Geological Sciences, The University of Texas at Austin, Austin, TX 78712, United States of America

## ARTICLE INFO

## Article history:

Received 20 August 2018

Received in revised form 19 January 2019

Accepted 21 January 2019

Available online 25 February 2019

## Keywords:

PDE-constrained optimal control

Taylor approximation

Variance reduction

Scalability

Uncertainty quantification

High dimensionality

## ABSTRACT

In this work we develop a scalable computational framework for the solution of PDE-constrained optimal control problems under high-dimensional uncertainty. Specifically, we consider a mean-variance formulation of the control objective and employ a Taylor expansion with respect to the uncertain parameter field either to directly approximate the control objective or as a control variate for variance reduction. The expressions for the mean and variance of the Taylor approximation are known analytically, although their evaluation requires efficient computation of the trace of the (preconditioned) Hessian of the control objective. We propose to estimate this trace by solving a generalized eigenvalue problem using a randomized algorithm that only requires the action of the Hessian on a small number of random directions. Then, the computational work does not depend on the nominal dimension of the uncertain parameter, but depends only on the effective dimension (i.e., the rank of the preconditioned Hessian), thus ensuring scalability to high-dimensional problems. Moreover, to increase the estimation accuracy of the mean and variance of the control objective by the Taylor approximation, we use it as a control variate for variance reduction, which results in considerable computational savings (several orders of magnitude) compared to a plain Monte Carlo method. In summary, our approach amounts to solving an optimal control problem constrained by the original PDE, the generalized eigenvalue equations at a small number of eigenfunctions, and a set of linearized PDEs that arise from the computation of the gradient and Hessian of the control objective with respect to the uncertain parameter. We use the Lagrangian formalism to derive expressions for the gradient with respect to the control and apply a gradient-based optimization method to solve the problem. We demonstrate the accuracy, efficiency, and scalability of the proposed computational method for two examples with high-dimensional uncertain parameters: subsurface flow in a porous medium modeled as an elliptic PDE, and turbulent jet flow modeled by the Reynolds-averaged Navier–Stokes equations coupled with a nonlinear advection–diffusion equation characterizing model uncertainty. In particular, for the latter more challenging example we show scalability of our algorithm up to one million parameters resulting from discretization of the uncertain parameter field.

© 2019 Elsevier Inc. All rights reserved.

\* Corresponding author.

E-mail address: peng@ices.utexas.edu (P. Chen).

## 1. Introduction

Optimal control (and more general optimization) constrained by partial differential equations (PDEs) are ubiquitous in many applications of practical relevance for science and engineering. The objective of such problems is to seek an optimal control that minimizes a cost functional, which often consists of a control objective and a penalty term. The control objective is often given as a distance between the solution of the PDE and a desired state or more generally formulated as a quantity of interest related to the solution, while the penalty term reflects the cost of the control or imposes regularity of the control. In practical applications, uncertainties in the PDE model are inevitable and can arise from various sources, such as the PDE coefficients, initial or boundary conditions, external sources, or geometries. Different realizations of these uncertain parameters may lead to significant differences in the optimal control. Mathematical theories and computational methods have been developed for several decades to deal with deterministic PDE-constrained optimal control [14,15,39,40,44,56,75]. More recently, PDE-constrained optimal control under uncertainty has become an important area of research and has drawn increasing attention [3,4,11,16,26–30,41,45,49–54,61,67,71,74].

Different formulations of the cost functional have been studied to incorporate uncertainty into the PDE-constrained optimal control. A straightforward choice is to optimize the mean of the control objective, i.e., the integration of the control objective with respect to the probability measure of the uncertain parameter [16,28,41,45,49]. However, this formulation does not control the variability of the control objective. To avoid a possible and undesired large variation of the control objective, which may represent the risk of a system failure, one can include the variance or higher moments of the control objective in the cost functional [11,67,74]. An alternative approach is use the Value-at-Risk (VaR) or conditional VaR [50], which measure the (conditional) probability of the control objective surpassing a certain critical value. An extreme choice is robust optimization (also referred as min-max optimization), where one seeks a control that yields the optimal value of the objective in the worst-case scenario, i.e., the extreme value of the control objective in the range of the uncertain parameter [13,54].

A common challenge in all of these formulations is to compute the statistics of the control objective (i.e., mean, variance, higher moments, or conditional probabilities); this involves the integration of the control objective or related quantities with respect to the probability measure of the uncertain parameter. This challenge becomes more severe for high-dimensional uncertain parameter space, which is often the case when the parameter is a uncertain spatially correlated field. A straightforward approach is to apply a Monte Carlo method to compute the statistics by taking the average of the control objective or related quantities at a set of samples randomly drawn according to the probability measure of the parameter. However, the converge rate of the Monte Carlo estimator is just  $O(M^{-1/2})$ , where  $M$  is the number of samples. This may lead to an extremely large number of samples to achieve a target accuracy, and therefore may not be feasible when the state equation is a complex nonlinear PDE. Quasi Monte Carlo methods using low-discrepancy sequences improve the convergence rate to  $O((\log(M))^d/M)$  for a  $d$ -dimensional parameter [34], and is thus attractive for low to medium dimensional problems. Stochastic Galerkin and stochastic collocation methods have also been used to compute the statistical moments in the cost functional [28,29,41,45,49,52–54,67,74], provided that a suitable finite dimensional parametrization of the uncertain parameter, such as a truncated Karhunen–Loève expansion, is applicable. These methods achieve fast convergence when the control objective depends smoothly on the low-dimensional parameter, but suffer from the so-called *curse of dimensionality*, i.e., the convergence rate quickly deteriorates as the dimension of the parameter increases. Recent advances in adaptive and anisotropic sparse quadrature [24,72] and high-order quasi Monte Carlo methods [35] have been shown to achieve a dimension-independent convergence rate  $N^{-s}$  with  $s$  potentially much larger than  $1/2$ . Here  $s$  is independent of the nominal dimension of the uncertain parameter and depends only on the choice of parametrization and on the regularity of the control objective with respect to the parameter, thus mitigating or even breaking the curse of dimensionality. However, the convergence may still be slow if the choice of the parametrization does not correctly capture the effective parameter dimensions.

Another computational challenge is that, in applications of practical relevance, the PDE may lead, after discretization, to possibly nonlinear large-scale systems that are extremely expensive to solve. Thus, only a limited number of high-fidelity PDE solves can be afforded. This challenge prevents a direct application of most of the computational methods introduced above as they require a large number of evaluations of the control objective for computing its statistics. To tackle this challenge, reduced basis methods [43,65,68] or other model reduction techniques [10] can be employed to exploit the low dimensionality of the solution manifold, even when the parameter lives in a high-dimensional space. Such methods solve the high-fidelity state equation at a few carefully selected samples of the parameter, and use such high-fidelity solutions (known as snapshots) as the basis to approximate the solution at any other realization of the uncertain parameter by a (Petrov)–Galerkin projection. Solving the projected problem in the reduced basis space allows for considerable computational saving compared to solving in the full space, such as a finite element space with a large number of degrees of freedom. These techniques have been successfully developed for solving parametric and stochastic PDE-constrained optimal control with several classes of PDE models [6,12,28,30,31,47,59,60,77]. Nevertheless, difficulties in using such techniques arise for highly nonlinear problems that require effective affine approximation (see [43,65] and references therein) or when the solution manifold becomes high-dimensional, even if the effective dimension of the manifold of the control objective is low [17,25].

In this work, we propose a scalable computational framework that uses Taylor expansions of the control objective with respect to the uncertain parameter to efficiently and accurately approximate the cost functional. This framework takes ad-

vantage both of the smooth dependence of the control objective on the uncertain parameter and of the low-rank structure of the Hessian of the control objective, which exposes the low effective dimension of the nominally high-dimensional parameter space. Such Taylor expansions can be used directly to approximate the moments of the control objective as proposed in [3] or as control variates to reduce the variance of Monte Carlo estimators. Analytic expressions for the mean and variance of the Taylor expansions are known; we refer to [3] for such expressions in the case of linear and quadratic approximations. These expressions depend on the trace of the preconditioned Hessian (second order derivative of the control objective with respect to the uncertain parameter) and/or its square. While for low-dimensional problems it is possible to explicitly compute the Hessian matrix (see [37,38]), this approach is not tractable for high-dimensional problems, where the Hessian is a formally large dense operator implicitly defined by its action on a given vector. Gaussian trace estimators (see [5]) only need to perform the action of the Hessian on a number of random directions, and have been successfully used for the solution of optimal experimental design in [1,2]. However, such estimators may still require a large number of Hessian actions to achieve high accuracy; see [5] for probabilistic bounds on the number of matrix actions necessary to achieve a target accuracy with high probability, even when the eigenvalues of the preconditioned Hessian decay very fast. In this work, we approximate the trace of the preconditioned Hessian with the sum of its dominant eigenvalues. Randomized algorithms [42,70] are computationally efficient methods to compute dominant eigenvalues of a large-scale implicitly defined operator whose spectra exhibit fast decay. The spectral properties of the Hessian operator have been extensively studied, and it has been observed numerically or proven analytically that for many problems the Hessian operator is either nearly low-rank or its eigenvalues exhibit fast decay [1–3,8,18–23,32,33,36,46,58,64]. Thus, this approach can provide highly accurate approximation of the trace at a cost (measured in number of Hessian actions) much smaller than the Gaussian trace estimators. We note that this approach results in an optimal control problem constrained by the generalized eigenvalue equations, with a small number of eigenfunctions computed at each optimization step. In contrast, a number of fixed eigenfunctions computed at some nominal value of the controls is used across the optimization iterations in [3]; as a result, the generalized eigenvalue equations are not considered as constraints for the optimal control problem. This latter approach in [3] likely works when the (subspace spanned by the dominant eigenfunctions of the) Hessian does not change much at different optimization iterations, but could fail to achieve the desired accuracy when the change is large.

The Taylor approximation may be not sufficiently accurate and introduces bias in the estimation of the mean and variance of the control objective. Specifically, the norm of the remainder of the Taylor expansion scales as  $O(\text{tr}(\mathcal{C})^{(r+1)/2})$  where  $\mathcal{C}$  is the covariance of the uncertain parameter and  $r$  is the order of the truncated Taylor expansion, e.g.,  $r = 1$  and  $2$  for linear and quadratic approximations, respectively [3]. To address this issue, we propose to use the Taylor approximation as a control variate to reduce the variance of the Monte Carlo estimator for the mean and variance of the control objective, or use Monte Carlo correction for the remainder of the Taylor expansion from another viewpoint. In cases when the Taylor approximation is highly correlated with the exact control objective, reduction in the computational cost can be dramatic, up to several orders of magnitude in our numerical examples.

We consider two examples of PDE-constrained optimal control under uncertainty to numerically illustrate the scalability, accuracy and efficiency of our computational framework. The first is a problem of optimal control of subsurface flow in a porous medium, where the control is the injection rate at given well locations, the state equation is an elliptic PDE representing single phase flow with uncertain permeability, and the objective is to drive the pressure field to a desired state. The second is an optimal boundary control for a turbulent jet flow, where the control is the inlet velocity profile, the state equations are the Reynolds-averaged Navier–Stokes equations coupled with a nonlinear stochastic advection–diffusion equation representing model uncertainty, and the objective is to maximize the jet width at a cross-section a certain distance from the inlet boundary. In both cases, the uncertain parameter is infinite-dimensional in the continuous setting and becomes high-dimensional after discretization.

For solution of the PDE-constrained optimal control with Taylor approximation and variance reduction we employ a gradient-based method, specifically the limited memory BFGS method [62]. The cost functional involves the analytical expressions for the mean and variance of the Taylor approximation and the Monte Carlo correction when using the Taylor approximation for variance reduction; the constraint involves the state (possibly nonlinear) PDE, the generalized eigenvalue equations, and a set of linearized PDEs used in the computation of the gradient and Hessian of the control objective with respect to the uncertain parameter. We use the Lagrangian formalism to derive the gradient of the cost functional with respect to the control. Specifically, we compute mixed high-order derivatives of the Lagrangian functional with respect to the state, the adjoint, the uncertain parameter, and the control by symbolic differentiation in a variational setting before discretization. This allows us to efficiently apply our framework to rather complex PDE models such as the turbulent jet example, which features coupled nonlinear PDEs, nonlinear numerical stabilization, and weak imposition of Dirichlet boundary conditions. We demonstrate numerically that our method scales independent of the uncertain parameter dimension, in that the number of required PDE solves does not depend on the discretization of the parameter field, and instead depends only on the effective or intrinsic dimension of the problem, i.e., the number of parameter dimensions to which the control objective is sensitive.

The rest of the paper is organized as follows: We present the basic setting for the PDE-constrained optimal control under uncertainty in Section 2 using a general notation for the uncertain parameter, the PDE model, the cost functional, and the optimal control problem. We also present a sample average approximation of the mean and variance of the control objective in this section. Section 3 is devoted to the development of the Taylor approximation and variance reduction for the evaluation of the mean and variance of the control objective. We also present computation of the gradient and Hessian

of the control objective with respect to the uncertain parameter and a double-pass randomized algorithm for solution of generalized eigenvalue problems. This is followed by Section 4 where we derive the gradient of the cost functional with respect to the control for different approximations including the sample average approximation, Taylor approximation, and Taylor approximation with Monte Carlo correction. Numerical results for two examples are presented in Section 5 for the demonstration of the scalability, accuracy, and efficiency of the proposed computational methods. Section 6 closes with some conclusions and perspectives.

## 2. PDE-constrained optimal control under uncertainty

In this section we present the PDE-constrained optimal control under uncertainty in an abstract setting. To this aim, let us first introduce the notation used throughout the paper. Let  $\mathcal{X}$  be a Hilbert space and  $\mathcal{X}'$  the dual space, then  $\mathcal{X} \langle \cdot, \cdot \rangle_{\mathcal{X}'}$  denotes the duality pairing between the spaces  $\mathcal{X}$  and  $\mathcal{X}'$ . For ease of notation, we will omit to specify the subscripts  $\mathcal{X}$  and  $\mathcal{X}'$  and simply write  $\langle \cdot, \cdot \rangle$  when the spaces can be inferred from the context without ambiguity. Given two Hilbert spaces  $\mathcal{X}$  and  $\mathcal{Y}$  and a map  $f : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}$ ,  $\partial_x f(x, y) \in \mathcal{X}'$  denotes the Fréchet derivative of  $f(x, y)$  with respect to  $x$  evaluated at  $(x, y)$ , which satisfies

$$\lim_{\tilde{x} \rightarrow 0} |f(x + \tilde{x}, y) - f(x, y) - \langle \tilde{x}, \partial_x f(x, y) \rangle| / \|\tilde{x}\|_{\mathcal{X}} = 0. \tag{1}$$

Let  $\partial_{xy} f(x, y) : \mathcal{Y} \mapsto \mathcal{X}'$  denote the Fréchet derivative of  $\partial_x f(x, y)$  with respect to  $y$  evaluated at  $(x, y)$ , or the second order (mixed) Fréchet derivative of  $f(x, y)$  with respect to  $x$  and  $y$  evaluated at  $(x, y)$ . Similarly,  $\partial_{yx} f(x, y) : \mathcal{X} \mapsto \mathcal{Y}'$  denotes the Fréchet derivative of  $\partial_y f(x, y)$  with respect to  $x$  evaluated at  $(x, y)$ , which is the adjoint operator of  $\partial_{xy} f(x, y)$  and satisfies

$$\langle \tilde{x}, \partial_{xy} f \hat{y} \rangle := \mathcal{X} \langle \tilde{x}, \partial_{xy} f \hat{y} \rangle_{\mathcal{X}'} = \mathcal{Y} \langle \hat{y}, \partial_{yx} f \tilde{x} \rangle_{\mathcal{Y}'} =: \langle \hat{y}, \partial_{yx} f \tilde{x} \rangle, \quad \forall \tilde{x} \in \mathcal{X}, \hat{y} \in \mathcal{Y}. \tag{2}$$

where we have omitted the argument  $(x, y)$  for simplicity. Finally, let us introduce the Hilbert spaces  $\mathcal{U}, \mathcal{V}, \mathcal{M}$ , and  $\mathcal{Z}$  in which the state  $u$ , the adjoint  $v$ , the uncertain parameter  $m$ , and the deterministic control  $z$  live, respectively. The exact definition of these spaces is problem specific and is specified in Sec. 5.

### 2.1. The uncertain parameter

We assume that the uncertain parameter  $m \in \mathcal{M}$  has mean  $\bar{m} \in \mathcal{M}$  and covariance  $\mathcal{C} : \mathcal{M} \rightarrow \mathcal{M}'$ . The computational framework developed in this work is applicable as long as this assumption is satisfied regarding the uncertain parameter, which can be both finite and infinite dimensional. In particular, to demonstrate the scalability of our computational methods, we consider an infinite-dimensional Gaussian random field with distribution  $\mu = \mathcal{N}(\bar{m}, \mathcal{C})$  for the uncertain parameter, which is one of the most important building blocks in modern spatial statistical modeling. A general covariance structure is known as Matérn covariance. It is equivalent to that the uncertain parameter  $m$  is given as the solution of some linear fractional stochastic PDE [55]. In this work we consider this type of covariance, i.e.,

$$\mathcal{C} = (-\alpha_1 \nabla \cdot (\Theta \nabla) + \alpha_2 I)^{-\alpha}, \tag{3}$$

where  $\nabla \cdot$ ,  $\nabla$ , and  $I$  are the divergence, gradient, and identity operators, respectively; the parameter  $\alpha > 0$  reflects the smoothness of the random field, which renders the covariance of trace class when  $\alpha > d/2$  for spatial dimension  $d = 1, 2, 3$ ,  $\alpha_1, \alpha_2 > 0$  controls its correlation length and variance, while  $\Theta \in \mathbb{R}^{d \times d}$  enables its spatial anisotropy.

### 2.2. The PDE model and control objective

We consider a state problem modeled by the following PDE model. Given a realization of the uncertain parameter  $m \in \mathcal{M}$  and a deterministic control  $z \in \mathcal{Z}$ , find the state  $u \in \mathcal{U}$ , such that

$$\mathcal{R}(u, m, z) = 0 \quad \text{in } \mathcal{V}', \tag{4}$$

where  $\mathcal{R}(\cdot, m, z) : \mathcal{U} \rightarrow \mathcal{V}'$  denotes a (possibly nonlinear) operator from  $\mathcal{U}$  to the dual of  $\mathcal{V}$ . We then define the weak form associated to the state equation (4) by means of duality pairing, and write

$$r(u, v, m, z) := \mathcal{V} \langle v, \mathcal{R}(u, m, z) \rangle_{\mathcal{V}'} = 0 \quad \forall v \in \mathcal{V}. \tag{5}$$

Note that  $r(u, v, m, z)$  is linear with respect to the adjoint  $v$  and possibly nonlinear with respect to the state  $u$ .

As for the control objective, we consider a real-valued, possibly nonlinear, functional  $Q$  of the state  $u \in \mathcal{U}$ ,

$$Q : \mathcal{U} \rightarrow \mathbb{R}. \tag{6}$$

For simplicity, we assume that  $Q$  depends on  $m$  and  $z$  only implicitly through the solution of the state equation (5). The general case where  $Q$  has an explicit dependence on  $m$  and  $z$  can be treated with slight modifications in the derivation of the cost functional and its gradient, as we address along the presentation in Sec. 3 and Sec. 4.

### 2.3. The PDE-constrained optimal control under uncertainty

Since the control objective  $Q$  (implicitly) depends on the uncertain parameter  $m$  through the state  $u$ , we face PDE-constrained optimal control under uncertainty. In this work we consider a mean-variance cost functional of the form

$$\mathcal{J}(z) = \mathbb{E}[Q] + \beta \text{Var}[Q] + \mathcal{P}(z), \tag{7}$$

which is often used in stochastic programming [73]. Above  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$  are, respectively, the mean and variance of  $Q$  with respect to the probability measure  $\mu$  of the uncertain parameter  $m$ , and are defined as

$$\mathbb{E}[Q] := \int_{\mathcal{M}} Q(u(m, z)) d\mu(m), \text{ and } \text{Var}[Q] := \int_{\mathcal{M}} (Q(u(m, z)) - \mathbb{E}[Q])^2 d\mu(m) = \mathbb{E}[Q^2] - (\mathbb{E}[Q])^2. \tag{8}$$

As the control objective  $Q$  (implicitly) depends on the control  $z$  through the state  $u$ , so do  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$ . The scaling factor  $\beta \geq 0$  in (7) controls the relative weight of the mean and variance, and  $\mathcal{P}(z)$  is a penalty term.

Recalling that the state  $u$  (implicitly) depends on the uncertain parameter  $m$  and control  $z$  through the solution of the PDE (5), we then formulate the PDE-constrained optimal control under uncertainty as: find  $z^* \in \mathcal{Z}$  such that

$$z^* = \arg \min_{z \in \mathcal{Z}} \mathcal{J}(z), \text{ subject to the problem (5)}. \tag{9}$$

To evaluate the mean and variance in the cost functional (7), a plain Monte Carlo sample average approximation can be used. We briefly discuss this approach below, while in the following section we introduce more sophisticated approximation that allow for more efficient algorithms by exploiting high-order derivatives.

### 2.4. Sample average approximation

The (plain Monte Carlo) sample average approximation of the mean reads

$$\mathbb{E}[Q] \approx \widehat{Q} := \frac{1}{M_1} \sum_{i=1}^{M_1} Q(m_i), \tag{10}$$

where  $m_i, i = 1, \dots, M_1$ , are independent identically distributed (i.i.d.) random samples drawn from the probability measure  $\mu$ . Similarly, the sample average approximation of the variance is given by

$$\text{Var}[Q] \approx \widehat{V}_Q := \frac{1}{M_2} \sum_{i=1}^{M_2} Q^2(m_i) - \left( \frac{1}{M_2} \sum_{i=1}^{M_2} Q(m_i) \right)^2, \tag{11}$$

where  $m_i, i = 1, \dots, M_2$ , are i.i.d. random samples from  $\mu$ . Since each evaluation of  $Q(m_i)$  involves solving the possibly nonlinear state equation (5), it is common practice to choose  $M_1 = M_2 = M$  and reuse the same samples  $m_i$  to estimate both the mean and variance of  $Q$ . However, a more careful choice of  $M_1$  and  $M_2$  relies on estimating the variance of  $Q$  and  $Q^2$  and aims at minimizing the mean square error of  $\widehat{Q} + \beta \widehat{V}_Q$ . Note that a large number of samples is required to obtain a relatively accurate approximation of  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$  due to the slow convergence of Monte Carlo sampling, thus rendering this approach computationally prohibitive for practical applications due to the large number of (nonlinear) PDE solves at each optimization step.

## 3. Taylor approximation and variance reduction

We develop the main approximation methods in this section, which include the approximation by (low-order) Taylor expansion of the control objective, the computation of the gradient and Hessian of the control objective with respect to the uncertain parameter used in the Taylor expansion, the randomized algorithm to solve a generalized eigenvalue problem for trace estimation, and the Monte Carlo correction for the remainder of the Taylor expansion.

### 3.1. Taylor approximation of the control objective

In this section we derive expressions for the Taylor expansion of the control objective  $Q(u(m, z))$  with respect to the uncertain parameter  $m$  for a fixed value of the control  $z$ . In what follows, we denote with the subscript  $m$  the total derivative of  $Q(u(m, z))$  with respect to  $m$  and we assume sufficient smoothness so that all Fréchet derivatives are well-defined. For ease of notation, we also denote with  $\langle \cdot, \cdot \rangle$  the duality-pairing  $\mathcal{M} \langle \cdot, \cdot \rangle_{\mathcal{M}'}$ , and, with a slight abuse of notation, we write  $Q(m)$  to indicate the control objective  $Q(u(m, z))$  evaluated at a fixed value of  $z$ .

A formal Taylor expansion of the control objective evaluated at  $\bar{m} \in \mathcal{M}$  is given in [3]

$$Q(m) = Q(\bar{m}) + \langle m - \bar{m}, Q_m(\bar{m}) \rangle + \frac{1}{2} \langle m - \bar{m}, Q_{mm}(\bar{m})(m - \bar{m}) \rangle + \dots, \quad (12)$$

where  $Q_m(\bar{m}) \in \mathcal{M}'$  and  $Q_{mm}(\bar{m}) : \mathcal{M} \rightarrow \mathcal{M}'$  are the first and second order Fréchet derivatives of  $Q$  with respect to  $m$  evaluated at the nominal  $\bar{m} \in \mathcal{M}$ . In what follows, we refer to such derivatives as the gradient and Hessian of  $Q$ , respectively. Based on the Taylor expansion, we can define the linear approximation of  $Q$  as

$$Q_{\text{lin}}(m) = Q(\bar{m}) + \langle m - \bar{m}, Q_m(\bar{m}) \rangle, \quad (13)$$

and the quadratic approximation of  $Q$  as

$$Q_{\text{quad}}(m) = Q(\bar{m}) + \langle m - \bar{m}, Q_m(\bar{m}) \rangle + \frac{1}{2} \langle m - \bar{m}, Q_{mm}(\bar{m})(m - \bar{m}) \rangle. \quad (14)$$

We remark that  $Q_{\text{lin}}(m)$  and  $Q_{\text{quad}}(m)$  are polynomials with respect to the uncertain parameter  $m$ , and depend on the control  $z$  through the solution of the state equation (5) at  $m = \bar{m}$ .

For the linear approximation, given the mean  $\bar{m}$  and covariance  $\mathcal{C}$  of the uncertain parameter  $m$ , we have

$$\mathbb{E}[Q_{\text{lin}}] = Q(\bar{m}), \text{ and } \text{Var}[Q_{\text{lin}}] = \langle Q_m(\bar{m}), \mathcal{C} Q_m(\bar{m}) \rangle. \quad (15)$$

In fact,  $Q_{\text{lin}}(m)$  is Gaussian if the uncertain parameter  $m \sim \mathcal{N}(\bar{m}, \mathcal{C})$ . For the quadratic approximation, we have

$$\mathbb{E}[Q_{\text{quad}}] = \mathbb{E}[Q_{\text{lin}}] + \frac{1}{2} \text{tr}(\mathcal{H}), \text{ and } \text{Var}[Q_{\text{quad}}] = \text{Var}[Q_{\text{lin}}] + \frac{1}{2} \text{tr}(\mathcal{H}^2), \quad (16)$$

where  $\mathcal{H} = \mathcal{C}^{1/2} Q_{mm}(\bar{m}) \mathcal{C}^{1/2}$  is the covariance preconditioned Hessian, and  $\text{tr}(\cdot)$  denotes the trace operator, see details in [3]. Efficient algorithms to accurately estimate the trace of  $\mathcal{H}$  and  $\mathcal{H}^2$  will be discussed in Sec. 3.3.

Finally, following [3], we approximate the cost functional  $J(z)$  in (7) by replacing  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$  with the moments of the linear or quadratic Taylor expansions of the control objective.

### 3.2. Computation of the gradient and Hessian of the control objective with respect to the uncertain parameter

In this section we use the Lagrangian formalism to derive the expressions for the gradient  $Q_m$  and for the action of the Hessian  $Q_{mm}$  on a given direction  $\hat{m}$  evaluated at  $\bar{m}$ . In what follows, for ease of notation, we use the overline to indicate quantities evaluated at  $m = \bar{m}$  and define

$$\bar{r} := r(u, v, \bar{m}, z), \text{ and } \bar{Q} := Q(u(\bar{m}, z)). \quad (17)$$

We define the Lagrangian functional

$$\mathcal{L}(u, v, m, z) = Q(u) + r(u, v, m, z), \quad (18)$$

where the adjoint  $v$  is regarded as the Lagrangian multiplier of the state equation.

By requiring the first order variation of (18) at  $\bar{m}$  with respect to the adjoint  $v$  to vanish, we obtain the state problem: find  $u \in \mathcal{U}$ , such that

$$\langle \tilde{v}, \partial_v \bar{r} \rangle = 0, \quad \forall \tilde{v} \in \mathcal{V}. \quad (19)$$

Similarly, by setting the first order variation of (18) at  $\bar{m}$  with respect to the state  $u$  to zero, we obtain the adjoint problem: find  $v \in \mathcal{V}$ , such that

$$\langle \tilde{u}, \partial_u \bar{r} \rangle = -\langle \tilde{u}, \partial_u \bar{Q} \rangle, \quad \forall \tilde{u} \in \mathcal{U}. \quad (20)$$

Then, the gradient of  $Q$  at  $\bar{m}$  acting in the direction  $\tilde{m} = m - \bar{m}$  is given by

$$\langle \tilde{m}, Q_m(\bar{m}) \rangle = \langle \tilde{m}, \partial_m \bar{r} \rangle, \quad (21)$$

where  $u$  solves the state problem (19) and  $v$  solves the adjoint problem (20).

Similarly, for the computation of the Hessian of  $Q$  acting on  $\hat{m}$ , we consider the Lagrangian functional

$$\mathcal{L}^H(u, v, m, z; \hat{u}, \hat{v}, \hat{m}) = \langle \hat{m}, \partial_m \bar{r} \rangle + \langle \hat{v}, \partial_v \bar{r} \rangle + \langle \hat{u}, \partial_u \bar{r} + \partial_u \bar{Q} \rangle, \quad (22)$$

where  $\hat{u}$  and  $\hat{v}$  denote the incremental state and incremental adjoint, respectively. By taking variation of (22) with respect to the adjoint  $v$  and using (2), we obtain the incremental state problem: find  $\hat{u} \in \mathcal{U}$  such that

$$\langle \tilde{v}, \partial_{vu} \bar{r} \hat{u} \rangle = -\langle \tilde{v}, \partial_{vm} \bar{r} \hat{m} \rangle, \quad \forall \tilde{v} \in \mathcal{V}, \tag{23}$$

where the derivatives  $\partial_{vu} \bar{r} : \mathcal{U} \rightarrow \mathcal{V}'$  and  $\partial_{vm} \bar{r} : \mathcal{M} \rightarrow \mathcal{V}'$  are linear operators. The incremental adjoint problem, obtained by taking variation of (22) with respect to the state  $u$  and using (2), reads: find  $\hat{v} \in \mathcal{V}$  such that

$$\langle \tilde{u}, \partial_{uv} \bar{r} \hat{v} \rangle = -\langle \tilde{u}, \partial_{uu} \bar{r} \hat{u} + \partial_{uu} \bar{Q} \hat{u} + \partial_{um} \bar{r} \hat{m} \rangle, \quad \forall \tilde{u} \in \mathcal{U}, \tag{24}$$

where  $\partial_{uv} \bar{r} : \mathcal{V} \rightarrow \mathcal{U}'$  is the adjoint of  $\partial_{vu} \bar{r} : \mathcal{U} \rightarrow \mathcal{V}'$  in the sense of (2). We remark that an extra term  $-\langle \tilde{u}, \partial_{um} \bar{Q} \hat{m} \rangle$  is added on the right hand side if  $Q$  explicitly depends on  $m$ . The Hessian of  $Q$  at  $\bar{m}$  acting on  $\hat{m}$  can then be computed by taking variation of (22) with respect to  $m$  and using (2) as

$$\langle \tilde{m}, Q_{mm}(\bar{m}) \hat{m} \rangle = \langle \tilde{m}, \partial_{mv} \bar{r} \hat{v} + \partial_{mu} \bar{r} \hat{u} + \partial_{mm} \bar{r} \hat{m} \rangle, \tag{25}$$

where the incremental state  $\hat{u}$  and adjoint  $\hat{v}$  solve (23) and (24), respectively. Note that  $\langle \tilde{m}, \partial_{mu} \bar{Q} \hat{u} + \partial_{mm} \bar{Q} \hat{m} \rangle$  is also added if  $Q$  explicitly depends on  $m$ .

We summarize the computation of the gradient and Hessian of the control objective  $Q$  with respect to the uncertain parameter  $m$  in Algorithm 1.

---

**Algorithm 1** Computation of the gradient and Hessian of  $Q$  with respect to  $m$ .

---

**Input:** control objective  $Q$ , PDE  $r(u, v, m, z)$ , mean  $\bar{m}$ , direction  $\hat{m}$ .

**Output:** the gradient  $Q_m(\bar{m})$  and Hessian  $Q_{mm}(\bar{m})$  acting on  $\hat{m}$ .

1. Solve the state problem (19) for  $u$ .
  2. Solve the adjoint problem (20) for  $v$ .
  3. Compute the gradient  $Q_m(\bar{m})$  by (21).
  4. Solve the incremental state problem (23) for  $\hat{u}$ .
  5. Solve the incremental adjoint problem (24) for  $\hat{v}$ .
  6. Compute the Hessian  $Q_{mm}(\bar{m})$  acting on  $\hat{m}$  by (25).
- 

### 3.3. Randomized algorithms to compute the trace

The first and second moments of the quadratic approximation  $Q_{\text{quad}}$  in equation (16) involve the trace of the covariance preconditioned Hessian  $\mathcal{H}$ .  $\mathcal{H}$  is formally a large scale dense operator and it is only implicitly defined through its action on a given direction as in (25).

The standard approach to estimate the trace of an implicitly defined operator is based on a Monte Carlo method, where one computes the action of the operator on  $N$  random directions sampled from a suitable distribution. For example, using the Gaussian trace estimator, one can approximate  $\text{tr}(\mathcal{H})$  and  $\text{tr}(\mathcal{H}^2)$  as

$$\text{tr}(\mathcal{H}) \approx \widehat{T}_1(\mathcal{H}) := \frac{1}{N} \sum_{j=1}^N \langle \hat{m}_j, Q_{mm}(\bar{m}) \hat{m}_j \rangle \text{ and } \text{tr}(\mathcal{H}^2) \approx \widehat{T}_1(\mathcal{H}^2) := \frac{1}{N} \sum_{j=1}^N \langle Q_{mm}(\bar{m}) \hat{m}_j, \mathcal{C} Q_{mm}(\bar{m}) \hat{m}_j \rangle, \tag{26}$$

where  $\hat{m}_j, j = 1, \dots, N$ , are *i.i.d.* random samples from the Gaussian distribution  $\mathcal{N}(0, \mathcal{C})$ . Lower bounds on the number of samples  $N$  to obtain a guaranteed probabilistic error bound for the trace are discussed in [5]. Such bounds hold for any symmetric positive defined operator, regardless of size and spectral properties. However the number of samples  $N$  necessary to obtain an accurate approximation may be prohibitively large and impractical: e.g., to guarantee a relative error of  $10^{-3}$  in the estimation of the trace more than  $10^6$  samples are need using the bounds in [5]. Improved bounds are demonstrated in [66] and show that the Gaussian estimator becomes very inefficient if the stable rank of the operator is small and that it allows for small  $N$  only if the eigenvalues are all of approximately the same size. This means that the randomized Gaussian estimator is not a viable solution to estimate the trace of  $\mathcal{H}$ , since it has been observed numerically or proven analytically that for many problems the Hessian operator is either nearly low-rank or its eigenvalues exhibit fast decay [1–3,8,18–23,32,33,36,46,58,64].

In this work, we propose to estimate the trace of  $\mathcal{H}$  and  $\mathcal{H}^2$  by means of a randomized approximated eigendecomposition. Specifically, we approximate  $\text{tr}(\mathcal{H})$  and  $\text{tr}(\mathcal{H}^2)$  by

$$\text{tr}(\mathcal{H}) \approx \widehat{T}_2(\mathcal{H}) := \sum_{j=1}^N \lambda_j(\mathcal{H}), \text{ and } \text{tr}(\mathcal{H}^2) \approx \widehat{T}_2(\mathcal{H}^2) := \sum_{j=1}^N \lambda_j^2(\mathcal{H}), \tag{27}$$

where  $\lambda_j(\mathcal{H}), j = 1, \dots, N$ , are the dominant eigenvalues of  $\mathcal{H}$  obtained by solving the generalized eigenvalue problem

$$\langle \phi, Q_{mm}(\bar{m}) \psi_j \rangle = \lambda_j \langle \phi, \mathcal{C}^{-1} \psi_j \rangle, \quad \forall \phi \in \mathcal{M}, \quad j = 1, \dots, N. \tag{28}$$

Above, the eigenvector  $\psi_j, j = 1, \dots, N$ , are  $\mathcal{C}^{-1}$ -orthonormal, that is

$$\langle \psi_j, C^{-1} \psi_i \rangle = \delta_{ij}, \quad i, j = 1, \dots, N, \tag{29}$$

where  $\delta_{ij}$  is the Kronecker delta. Note that  $N$  is independent of the nominal parameter dimension, and, as shown numerically in Sec. 5, it is relatively small for  $\mathcal{H}$  nearly low-rank or when its spectrum exhibits fast decay.

After suitable discretization, e.g., by a finite element discretization, the generalized eigenvalue problem (28) results in an algebraic eigenproblem of the form  $A\psi = \lambda B\psi$  with  $A, B \in \mathbb{R}^{n \times n}$  and  $\psi \in \mathbb{R}^n$ . Algorithm 2 summarizes the so-called *double pass randomized algorithm* to solve the algebraic generalized eigenvalue problem (see [42,70] for details of the algorithms and [76] for its implementation).

---

**Algorithm 2** Randomized algorithm for generalized eigenvalue problem  $(A, B)$ .

---

**Input:** matrix  $A, B$ , the number of eigenpairs  $k$ , an oversampling factor  $p$ .

**Output:** eigenpairs  $(\Lambda_k, \Psi_k)$  with  $\Lambda_k = \text{diag}(\lambda_1, \dots, \lambda_k)$  and  $\Psi_k = (\psi_1, \dots, \psi_k)$ .

1. Draw a Gaussian random matrix  $\Omega \in \mathbb{R}^{n \times (k+p)}$ .
  2. Compute  $Y = B^{-1}(A\Omega)$ .
  3. Compute QR-factorization  $Y = QR$  such that  $Q^T BQ = I_{k+p}$ .
  4. Form  $T = Q^T A Q$  and compute eigendecomposition  $T = S\Lambda S^T$ .
  5. Extract  $\Lambda_k = \Lambda(1:k, 1:k)$  and  $\Psi_k = QS_k$  with  $S_k = S(:, 1:k)$ .
- 

There are multiple advantages in using Algorithm 2 to approximate the trace, see e.g., [42,69]. In terms of accuracy, the approximation error is bounded by the sum of the remaining eigenvalues, so that the error is small if the eigenvalues decay fast or if the Hessian  $\mathcal{H}$  has low rank, see [42,69] for more details. In terms of computational efficiency, the  $2(k+p)$  Hessian matrix-vector products, which entail the solution of a pair of linearized state/adjoint equations (as shown in Sec. 3.2) and therefore dominate the computational cost of the algorithm, can be computed independently and, therefore, asynchronously across the random directions.

### 3.4. Monte Carlo correction

Using the moments of the truncated Taylor expansion  $Q_{\text{lin}}$  and  $Q_{\text{quad}}$  as approximation of  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$  in the cost functional (7) may be not sufficiently accurate. In this circumstance, we can use the Taylor approximation as a control variate to reduce the variance of the Monte Carlo estimator. The basic idea is to correct the moments of  $Q_{\text{lin}}$  and  $Q_{\text{quad}}$  by applying a Monte Carlo method to estimate the mean (and variance) of the difference between  $Q$  and its Taylor approximation.

Specifically, recalling (15) and (13), the Monte Carlo correction for the mean of the linear approximation reads

$$\mathbb{E}[Q] = \mathbb{E}[Q_{\text{lin}}] + \mathbb{E}[Q - Q_{\text{lin}}] \approx \widehat{Q}_{\text{lin}} := Q(\bar{m}) + \frac{1}{M_1} \sum_{i=1}^{M_1} \left( Q(m_i) - Q(\bar{m}) - \langle m_i - \bar{m}, Q_m(\bar{m}) \rangle \right). \tag{30}$$

Similarly, recalling that the variance is invariant with respect to deterministic translations, we have

$$\begin{aligned} \text{Var}[Q] &= \text{Var}[Q - Q(\bar{m})] = \mathbb{E}[(Q - Q(\bar{m}))^2] - (\mathbb{E}[Q - Q(\bar{m})])^2 \\ &= \mathbb{E}[(Q_{\text{lin}} - Q(\bar{m}))^2] + \mathbb{E}[(Q - Q(\bar{m}))^2 - (Q_{\text{lin}} - Q(\bar{m}))^2] \\ &\quad - (\mathbb{E}[Q_{\text{lin}} - Q(\bar{m})] + \mathbb{E}[(Q - Q(\bar{m})) - (Q_{\text{lin}} - Q(\bar{m}))])^2, \end{aligned} \tag{31}$$

and therefore the Monte Carlo correction of the variance of the linear approximation is given by

$$\begin{aligned} \widehat{V}_Q^{\text{lin}} &:= \langle C Q_m(\bar{m}), Q_m(\bar{m}) \rangle + \frac{1}{M_2} \sum_{i=1}^{M_2} \left( (Q(m_i) - Q(\bar{m}))^2 - (\langle m_i - \bar{m}, Q_m(\bar{m}) \rangle)^2 \right) \\ &\quad - \left( \frac{1}{M_2} \sum_{i=1}^{M_2} Q(m_i) - Q(\bar{m}) - \langle m_i - \bar{m}, Q_m(\bar{m}) \rangle \right)^2. \end{aligned} \tag{32}$$

Note that, similar to the sample average approximation in Sec. 2.4, for simplicity we can use the same  $M = M_1 = M_2$  random samples for the approximation of both the mean and variance. A more careful choice of  $M_1$  and  $M_2$  relies on the variance of  $Q^{(1)} = Q - Q_{\text{lin}}$  and  $Q^{(2)} = (Q - \bar{Q})^2 - (Q - Q_{\text{lin}})^2$ , in order to balance the errors in the approximation of the mean and variance. For the quadratic approximation, a combination of (16) and (14) leads to

$$\begin{aligned} \mathbb{E}[Q] &= \mathbb{E}[Q_{\text{quad}}] + \mathbb{E}[Q - Q_{\text{quad}}] \approx \widehat{Q}_{\text{quad}} := Q(\bar{m}) + \frac{1}{2} \text{tr}(\mathcal{H}) \\ &\quad + \frac{1}{M_1} \sum_{i=1}^{M_1} \left( Q(m_i) - Q(\bar{m}) - \langle m_i - \bar{m}, Q_m(\bar{m}) \rangle - \frac{1}{2} \langle m_i - \bar{m}, Q_{mm}(\bar{m})(m_i - \bar{m}) \rangle \right). \end{aligned} \tag{33}$$

For the variance, akin to the linear approximation, we obtain

$$\begin{aligned} \text{Var}[Q] &= \mathbb{E}[(Q_{\text{quad}} - Q(\bar{m}))^2] + \mathbb{E}[(Q - Q(\bar{m}))^2 - (Q_{\text{quad}} - Q(\bar{m}))^2] \\ &\quad - (\mathbb{E}[Q_{\text{quad}} - Q(\bar{m})] + \mathbb{E}[(Q - Q(\bar{m})) - (Q_{\text{quad}} - Q(\bar{m}))])^2, \end{aligned} \tag{34}$$

which can be approximated by

$$\begin{aligned} \widehat{V}_Q^{\text{quad}} &:= \langle \mathcal{C} Q_m(\bar{m}), Q_m(\bar{m}) \rangle + \frac{1}{4}(\text{tr}(\mathcal{H}))^2 + \frac{1}{2}\text{tr}(\mathcal{H}^2) \\ &\quad + \frac{1}{M_2} \sum_{i=1}^{M_2} \left( (Q(m_i) - Q(\bar{m}))^2 - \left( \langle m_i - \bar{m}, Q_m(\bar{m}) \rangle + \frac{1}{2} \langle m_i - \bar{m}, Q_{mm}(\bar{m})(m_i - \bar{m}) \rangle \right)^2 \right) \\ &\quad - \left( \frac{1}{2} \text{tr}(\mathcal{H}) + \frac{1}{M_2} \sum_{i=1}^{M_2} \left( Q(m_i) - \langle m_i - \bar{m}, Q_m(\bar{m}) \rangle - \frac{1}{2} \langle m_i - \bar{m}, Q_{mm}(\bar{m})(m_i - \bar{m}) \rangle \right) \right)^2. \end{aligned} \tag{35}$$

Note that to achieve certain approximation accuracy, the number of Monte Carlo samples depends on the variance of the integrand. If the approximations  $Q_{\text{lin}}$  and  $Q_{\text{quad}}$  are highly correlated with  $Q$ , the variances of  $Q - Q_{\text{lin}}$  and  $Q - Q_{\text{quad}}$  would be small so that only a small number of Monte Carlo samples are required. This use of the Taylor approximation (of the control objective with respect to the uncertain parameter) can be viewed as a kind of “multifidelity Monte Carlo method” [63], in which the Taylor approximation plays the role of a reduced model.

#### 4. Gradient-based optimization

In this section, we present a gradient-based optimization method to solve the problem (9) using different approximations introduced in Sec. 3, namely the sample average approximation, the first and second order Taylor approximation, and variance reduction using the Taylor approximation as the control variate. These approximated formulations feature different levels of accuracy and computational complexity, and could possibly be combined within a multifidelity framework to obtain an efficient computational procedure while preserving high accuracy. In this section, we use the Lagrangian formalism to derive expressions for the gradient of the approximated formulations of the cost functional with respect to the control  $z$ , which are needed to perform the gradient-based optimization. In the following, to distinguish the gradient of the cost function with respect to the control from the gradient of the control objective with respect to the uncertain parameter, we denote the former as  $z$ -gradient.

##### 4.1. $z$ -gradient for the sample average approximation

By the sample average approximation presented in Sec. 2.4, the cost functional becomes

$$\mathcal{J}^{\text{MC}}(z) = \widehat{Q} + \beta \widehat{V}_Q + \mathcal{P}(z), \tag{36}$$

subject to the state problem (5) at  $m_i$ ,  $i = 1, \dots, M$ , where  $M = M_1 = M_2$  if the same samples are used for the mean and variance or  $M = M_1 + M_2$  otherwise. To compute the gradient of (36) with respect to the control  $z$  ( $z$ -gradient for short), we consider the Lagrangian functional

$$\mathcal{L}^{\text{MC}}(\{u_i\}, \{v_i\}, z) = \mathcal{J}^{\text{MC}}(z) + \sum_{i=1}^M r(u_i, v_i, m_i, z), \tag{37}$$

where  $\{v_i\}$  are the Lagrange multipliers (adjoints) of the state problems: given  $m_i \in \mathcal{M}$  and  $z \in \mathcal{Z}$ , find  $u_i \in \mathcal{U}$  such that

$$r(u_i, v_i, m_i, z) = 0, \quad \forall v_i \in \mathcal{V}; \quad i = 1, \dots, M. \tag{38}$$

The  $i$ -th adjoint problem ( $i = 1, \dots, M$ ) is obtained by setting to zero the first variation of the Lagrangian functional with respect to the state  $u_i$ . The  $i$ -th adjoint  $v_i$  then solves: given  $m_i \in \mathcal{M}$ ,  $z \in \mathcal{Z}$ , and  $u_i \in \mathcal{U}$  as the solution of (38), find  $v_i \in \mathcal{V}$  such that

$$\begin{aligned} \langle \tilde{u}_i, \partial_u r(u_i, v_i, m_i, z) \rangle &= -\frac{1}{M_1} \langle \tilde{u}_i, \partial_u Q(m_i) \rangle - \frac{2\beta}{M_2} Q(m_i) \langle \tilde{u}_i, \partial_u Q(m_i) \rangle \\ &\quad + \frac{2\beta}{M_2} \left( \frac{1}{M_2} \sum_{j=1}^{M_2} Q(m_j) \right) \langle \tilde{u}_i, \partial_u Q(m_i) \rangle, \quad \forall \tilde{u}_i \in \mathcal{U}. \end{aligned} \tag{39}$$

Then we can compute the  $z$ -gradient as

$$\langle \tilde{z}, D_z \mathcal{J}^{\text{MC}}(z) \rangle = \langle \tilde{z}, \partial_z \mathcal{L}^{\text{MC}}(\{u_i\}, \{v_i\}, z) \rangle = \langle \tilde{z}, D_z \mathcal{P}(z) \rangle + \sum_{i=1}^M \langle \tilde{z}, \partial_z r(u_i, v_i, m_i, z) \rangle, \tag{40}$$

where  $\{u_i\}$  and  $\{v_i\}$  denote the set of solutions of the state problem (38) and adjoint problem (39) at the set of samples  $\{m_j\}$ , respectively. Therefore, each gradient evaluation requires  $M$  (nonlinear) state PDE solves and  $M$  linearized PDE solves. Note that, for the general case in which  $Q$  explicitly depends on  $z$ , extra terms involving  $\langle \tilde{z}, \partial_z Q(m_i) \rangle$  need to be included in the gradient computation.

#### 4.2. z-gradient for the Taylor approximation

We first consider the linear approximation (15), by which the cost functional (7) becomes

$$\mathcal{J}_{\text{lin}}(z) = \bar{Q} + \beta \langle C Q_m(\bar{m}), Q_m(\bar{m}) \rangle + \mathcal{P}(z), \tag{41}$$

where  $\bar{Q}$  and its gradient  $Q_m(\bar{m})$  are defined in (17) and (21), respectively.

To compute the z-gradient of (41), we introduce the Lagrangian functional

$$\mathcal{L}_{\text{lin}}(u, v, u^*, v^*, z) = \mathcal{J}_{\text{lin}}(z) + \langle v^*, \partial_v \bar{r} \rangle + \langle u^*, \partial_u \bar{r} + \partial_u \bar{Q} \rangle, \tag{42}$$

where the adjoints  $v^* \in \mathcal{V}$  and  $u^* \in \mathcal{U}$  represent the Lagrange multipliers of the state problem (19) and adjoint problem (20).

By setting to zero the first order variation of the Lagrangian (42) with respect to  $v$  and using (2), we have: find  $u^* \in \mathcal{U}$  such that

$$\langle \tilde{v}, \partial_{vu} \bar{r} u^* \rangle = -2\beta \langle \tilde{v}, \partial_{vm} \bar{r} (C \partial_m \bar{r}) \rangle, \quad \forall \tilde{v} \in \mathcal{V}, \tag{43}$$

while setting to zero the first order variation with respect to  $u$  and using (2), we have: find  $v^* \in \mathcal{V}$  such that

$$\langle \tilde{u}, \partial_{uv} \bar{r} v^* \rangle = - \langle \tilde{u}, \partial_u \bar{Q} \rangle - 2\beta \langle \tilde{u}, \partial_{um} \bar{r} (C \partial_m \bar{r}) \rangle - \langle \tilde{u}, \partial_{uu} \bar{r} u^* + \partial_{uu} \bar{Q} u^* \rangle, \quad \forall \tilde{u} \in \mathcal{U}. \tag{44}$$

Finally, the z-gradient of  $\mathcal{J}_{\text{lin}}$  in a direction  $\tilde{z} \in \mathcal{Z}$  is computed as

$$\langle \tilde{z}, D_z \mathcal{J}_{\text{lin}}(z) \rangle = \langle \tilde{z}, \partial_z \mathcal{L}_{\text{lin}}(u, v, u^*, v^*, z) \rangle = 2\beta \langle \tilde{z}, \partial_{zm} \bar{r} (C \partial_m \bar{r}) \rangle + \langle \tilde{z}, \partial_z \mathcal{P}(z) \rangle + \langle \tilde{z}, \partial_{zv} \bar{r} v^* + \partial_{zu} \bar{r} u^* \rangle. \tag{45}$$

In summary, evaluation of the cost functional with the linear approximation requires solution of the state problem (19) and adjoint problem (20). Computation of the z-gradient requires, in addition, solution of the two linear problems in (43) and (44).

For the quadratic approximation (14), thanks to the relation (16) and the approximation (26), we obtain the approximation of the cost functional as

$$\mathcal{J}_{\text{quad}}(z) = \bar{Q} + \frac{1}{2} \sum_{j=1}^N \lambda_j + \beta \left( \langle C Q_m(\bar{m}), Q_m(\bar{m}) \rangle + \frac{1}{2} \sum_{j=1}^N \lambda_j^2 \right) + \mathcal{P}(z), \tag{46}$$

which is subject to the state and adjoint problems (19) and (20) for the evaluation of  $Q$  and its gradient at  $\bar{m}$ , the generalized eigenvalue problems (28) for the computation of the eigenvalues, as well as the incremental state and adjoint problems (23) and (24) for the Hessian action in (28). Correspondingly, we form the Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{quad}} & \left( u, v, \{\lambda_j\}, \{\psi_j\}, \{\hat{u}_j\}, \{\hat{v}_j\}, u^*, v^*, \{\lambda_j^*\}, \{\psi_j^*\}, \{\hat{u}_j^*\}, \{\hat{v}_j^*\}, z \right) \\ & = \mathcal{J}_{\text{quad}}(z) + \langle v^*, \partial_v \bar{r} \rangle + \langle u^*, \partial_u \bar{r} + \partial_u \bar{Q} \rangle + \sum_{j=1}^N \langle \psi_j^*, (Q_{mm}(\bar{m}) - \lambda_j C^{-1}) \psi_j \rangle + \sum_{j=1}^N \lambda_j^* \langle \langle \psi_j, C^{-1} \psi_j \rangle - 1 \rangle \\ & + \sum_{j=1}^N \langle \hat{v}_j^*, \partial_{vu} \bar{r} \hat{u}_j + \partial_{vm} \bar{r} \psi_j \rangle + \sum_{j=1}^N \langle \hat{u}_j^*, \partial_{uv} \bar{r} \hat{v}_j + \partial_{uu} \bar{r} \hat{u}_j + \partial_{uu} \bar{Q} \hat{u}_j + \partial_{um} \bar{r} \psi_j \rangle. \end{aligned} \tag{47}$$

Here we assume that the dominating eigenvalues are not repeated (thanks to their rapid decay as shown in Sec. 5), so that the constraints  $\langle \psi_j, C^{-1} \psi_j \rangle = 1, j = 1, \dots, N$ , are sufficient to guarantee the orthonormality  $\langle \psi_j, C^{-1} \psi_i \rangle = \delta_{ij}$  for any  $i, j = 1, \dots, N$ . In fact, for any  $i \neq j$ , by definition we have

$$\langle \psi_j, Q_{mm} \psi_i \rangle = \langle \psi_j, \lambda_i C^{-1} \psi_i \rangle \text{ and } \langle \psi_i, Q_{mm} \psi_j \rangle = \langle \psi_i, \lambda_j C^{-1} \psi_j \rangle, \tag{48}$$

which, by the symmetry of  $Q_{mm}$  and  $C^{-1}$ , leads to

$$(\lambda_i - \lambda_j)\langle \psi_i, C^{-1}\psi_j \rangle = 0, \text{ so that } \langle \psi_i, C^{-1}\psi_j \rangle = 0 \text{ if } \lambda_i \neq \lambda_j. \tag{49}$$

By setting the variation of  $\mathcal{L}_{\text{quad}}$  with respect to  $\lambda_j$  as zero, we obtain

$$\psi_j^* = \frac{1 + 2\beta\lambda_j}{2}\psi_j, \quad j = 1, \dots, N. \tag{50}$$

Subsequently, for each  $j = 1, \dots, N$ , taking the variation of  $\mathcal{L}_{\text{quad}}$  with respect to  $\hat{v}_j$  as zero, and using the Hessian action equation (25) and (2), we have: find  $\hat{u}_j^* \in \mathcal{U}$  such that

$$\langle \tilde{v}, \partial_{vu}\bar{r}\hat{u}_j^* \rangle = -\langle \tilde{v}, \partial_{vm}\bar{r}\psi_j^* \rangle, \quad \forall \tilde{v} \in \mathcal{V}, \tag{51}$$

which is the same as the incremental forward problem (23). Therefore, from the result (50), we have

$$\hat{u}_j^* = \frac{1 + 2\beta\lambda_j}{2}\hat{u}_j, \quad j = 1, \dots, N. \tag{52}$$

Then, taking variation with respect to  $\hat{u}_j$  as zero and using (2), we have: find  $\hat{v}_j^* \in \mathcal{V}$  such that

$$\langle \tilde{u}, \partial_{uv}\bar{r}\hat{v}_j^* \rangle = -\langle \tilde{u}, \partial_{uu}\bar{r}\hat{u}_j^* + \partial_{uu}\bar{Q}\hat{u}_j^* + \partial_{um}\bar{r}\psi_j^* \rangle, \quad \forall \tilde{u} \in \mathcal{U}, \tag{53}$$

which is the same as the incremental adjoint problem (24). Therefore, by (52) and (50), we obtain

$$\hat{v}_j^* = \frac{1 + 2\beta\lambda_j}{2}\hat{v}_j, \quad j = 1, \dots, N. \tag{54}$$

By setting to zero the variation with respect to  $v$  and using (2), we have: find  $u^* \in \mathcal{U}$  such that

$$\begin{aligned} \langle \tilde{v}, \partial_{vu}\bar{r}u^* \rangle &= -2\beta\langle \tilde{v}, \partial_{vm}\bar{r}(C\partial_m\bar{r}) \rangle - \sum_{j=1}^N \langle \tilde{v}, \partial_{vmu}\bar{r}\hat{u}_j\psi_j^* + \partial_{vmm}\bar{r}\psi_j\psi_j^* \rangle \\ &\quad - \sum_{j=1}^N \langle \tilde{v}, \partial_{vu}\bar{r}\hat{u}_j\hat{u}_j^* + \partial_{vum}\bar{r}\psi_j\hat{u}_j^* \rangle, \quad \forall \tilde{v} \in \mathcal{V}. \end{aligned} \tag{55}$$

Lastly, by setting to zero the variation with respect to  $u$  and using (2), we have: find  $v^* \in \mathcal{V}$  such that

$$\begin{aligned} \langle \tilde{u}, \partial_{uv}\bar{r}v^* \rangle &= -\langle \tilde{u}, \partial_u\bar{Q} \rangle - 2\beta\langle \tilde{u}, \partial_{um}\bar{r}(C\partial_m\bar{r}) \rangle - \langle \tilde{u}, \partial_{uu}\bar{r}u^* + \partial_{uu}\bar{Q}u^* \rangle \\ &\quad - \sum_{j=1}^N \langle \tilde{u}, \partial_{umv}\bar{r}\hat{v}_j\psi_j^* + \partial_{umu}\bar{r}\hat{u}_j\psi_j^* + \partial_{uum}\bar{r}\psi_j\psi_j^* \rangle - \sum_{j=1}^N \langle \tilde{u}, \partial_{uvu}\bar{r}\hat{u}_j\hat{v}_j^* + \partial_{uvm}\bar{r}\psi_j\hat{v}_j^* \rangle \\ &\quad - \sum_{j=1}^N \langle \tilde{u}, \partial_{uu}\bar{r}\hat{v}_j\hat{u}_j^* + \partial_{uuu}\bar{r}\hat{u}_j\hat{u}_j^* + \partial_{uuu}\bar{Q}\hat{u}_j\hat{u}_j^* + \partial_{uum}\bar{r}\psi_j\hat{u}_j^* \rangle, \quad \forall \tilde{u} \in \mathcal{U}, \end{aligned} \tag{56}$$

where the third order derivatives take the form of the bilinear map  $\partial_{xyz}f : \mathcal{Z} \times \mathcal{Y} \rightarrow \mathcal{X}'$ . Then, the  $z$ -gradient of the cost functional in the direction  $\tilde{z} \in \mathcal{Z}$  can be computed as

$$\langle \tilde{z}, D_z\mathcal{J}_{\text{quad}}(z) \rangle = \langle \tilde{z}, \partial_z\mathcal{L}_{\text{quad}}(\text{states, adjoints}, z) \rangle, \tag{57}$$

where states and adjoints represent the state and adjoint variables for short. Note that we do not need to solve for  $\lambda_j^*$  as the terms involving  $\lambda_j^*$  are independent of  $z$ . In summary, to get these states and adjoints, we need to solve 1 (nonlinear) state problem and  $2N + 3$  linear problems (1 for  $v$ ,  $N$  for  $\{\hat{u}_j\}$ ,  $N$  for  $\{\hat{v}_j\}$ , 1 for  $u^*$ , and 1 for  $v^*$ ), where the linear operators are the same as those in (23) and (24). Moreover, to compute the eigenvalues by Algorithm 2, which requires the action of the Hessian on  $N + p$  random directions twice, we need to solve  $4(N + p)$  linear problems.

We summarize the evaluation of the cost functional and its  $z$ -gradient with quadratic approximation as follows.

### 4.3. $z$ -gradient for the Monte Carlo correction

The approximation in (41) and (46) introduced by the Taylor approximation may be not sufficiently accurate and generates bias in the evaluation of the mean  $\mathbb{E}[Q]$  and variance  $\text{Var}[Q]$ . Recalling (30) and (32), we obtain a MC-corrected unbiased linear approximation for the cost functional, which is given by

$$\mathcal{J}_{\text{lin}}^{\text{MC}}(z) = \widehat{Q}_{\text{lin}} + \beta\widehat{V}_Q^{\text{lin}} + \mathcal{P}(z), \tag{58}$$

**Algorithm 3** Computation of the cost functional  $\mathcal{J}_{\text{quad}}$  and its  $z$ -gradient.

- Input:** cost functional  $\mathcal{J}(z)$  and PDE  $r(u, v, m, z)$ .  
**Output:** the approximate cost functional  $\mathcal{J}_{\text{quad}}$  and its  $z$ -gradient  $D_z \mathcal{J}_{\text{quad}}$ .
1. Solve the state problem (19) for  $u$  at  $\bar{m}$  and evaluate  $\bar{Q} = Q(u(\bar{m}))$ .
  2. Solve the adjoint problem (20) for  $v$  and evaluate  $Q_m(\bar{m})$  by (21).
  3. Use Algorithm 1 and 2 to compute the generalized eigenpairs  $(\lambda_j, \psi_j)_{j=1}^N$ .
  4. Compute the approximate cost functional  $\mathcal{J}_{\text{quad}}$  by (46).
  5. Solve the linear problem (55) for  $\hat{u}^*$ .
  6. Solve the linear problem (56) for  $\hat{v}^*$ .
  7. Compute the  $z$ -gradient  $D_z \mathcal{J}_{\text{quad}}$  by (57).

subject to the state problem (19) at  $\bar{m}$ , the adjoint problem (20), and the state problems (5) at  $m_i, i = 1, \dots, M$ , where  $M = M_1 = M_2$  if the same samples are used for the mean and variance or  $M = M_1 + M_2$  otherwise. We form the corresponding Lagrangian functional as

$$\mathcal{L}_{\text{lin}}^{\text{MC}}(u, v, \{u_i\}, u^*, v^*, \{v_i\}, z) = \mathcal{J}_{\text{lin}}^{\text{MC}}(z) + \langle v^*, \partial_v \bar{r} \rangle + \langle u^*, \partial_u \bar{r} + \partial_u \bar{Q} \rangle + \sum_{i=1}^M r(u_i, v_i, m_i, z). \tag{59}$$

Setting variation with respect to  $v^*, v_i$ , and  $u^*$  to zero we obtain  $u, u_i$ , and  $v$  by solving  $1 + M$  state problems (5) at  $m_i, i = 0, \dots, M$ , and 1 linear adjoint problem (20). In a similar way, setting variation with respect to  $v, u_i$ , and  $u$  to zero, we obtain  $u^*, v_i$ , and  $v^*$ , by solving  $2 + M$  linear problems. Thus, in addition,  $2 + M$  linear problems need to be solved to compute the  $z$ -gradient

$$\langle \bar{z}, D_z \mathcal{J}_{\text{lin}}^{\text{MC}}(z) \rangle = \langle \bar{z}, \partial_z \mathcal{L}_{\text{lin}}^{\text{MC}}(u, v, \{u_i\}, u^*, v^*, \{v_i\}, z) \rangle. \tag{60}$$

As for the quadratic approximation (14), by the Monte Carlo correction (33) and (35), we have an unbiased and more accurate evaluation of the cost functional as

$$\mathcal{J}_{\text{quad}}^{\text{MC}}(z) = \hat{Q}_{\text{quad}} + \beta \hat{V}_Q^{\text{quad}} + \mathcal{P}(z), \tag{61}$$

subject to the state problems (5) and (19), the adjoint problem (20), the incremental state and adjoint problems (23) and (24). We form the Lagrangian with the Lagrange multiplier  $(u^*, v^*, \{v_i\}, \{\lambda_j^*\}, \{\psi_j^*\}, \{\hat{u}_j^*\}, \{\hat{v}_j^*\}, \{\hat{u}_i^*\}, \{\hat{v}_i^*\})$  as

$$\begin{aligned} \mathcal{L}_{\text{quad}}^{\text{MC}}(u, v, \{u_i\}, \{\lambda_j\}, \{\psi_j\}, \{\hat{u}_j\}, \{\hat{v}_j\}, \{\hat{u}_i\}, \{\hat{v}_i\}, u^*, v^*, \{v_i\}, \{\lambda_j^*\}, \{\psi_j^*\}, \{\hat{u}_j^*\}, \{\hat{v}_j^*\}, \{\hat{u}_i^*\}, \{\hat{v}_i^*\}) \\ = \mathcal{J}_{\text{quad}}^{\text{MC}} + \langle v^*, \partial_v \bar{r} \rangle + \langle u^*, \partial_u \bar{r} + \partial_u \bar{Q} \rangle + \sum_{i=1}^M r(u_i, v_i, m_i, z) \\ + \sum_{j=1}^N \langle \psi_j^*, (Q_{mm}(\bar{m}) - \lambda_j C^{-1}) \psi_j \rangle + \sum_{j=1}^N \lambda_j^* (\langle \psi_j, C^{-1} \psi_j \rangle - 1) \\ + \sum_{j=1}^N \langle \hat{v}_j^*, \partial_{vu} \bar{r} \hat{u}_j + \partial_{vm} \bar{r} \psi_j \rangle + \sum_{j=1}^N \langle \hat{u}_j^*, \partial_{uv} \bar{r} \hat{v}_j + \partial_{uu} \bar{r} \hat{u}_j + \partial_{uu} \bar{Q} \hat{u}_j + \partial_{um} \bar{r} \psi_j \rangle \\ + \sum_{i=1}^M \langle \hat{v}_i^*, \partial_{vu} \bar{r} \hat{u}_i + \partial_{vm} \bar{r} m_i \rangle + \sum_{i=1}^M \langle \hat{u}_i^*, \partial_{uv} \bar{r} \hat{v}_i + \partial_{uu} \bar{r} \hat{u}_i + \partial_{uu} \bar{Q} \hat{u}_i + \partial_{um} \bar{r} m_i \rangle. \end{aligned} \tag{62}$$

By setting variation with respect to  $v^*, u^*, v_i, \hat{v}_j^*, \hat{u}_j^*, \hat{v}_i^*, \hat{u}_i^*$  to zero, we obtain the states  $u, v, u_i, \hat{u}_j, \hat{v}_j, \hat{u}_i, \hat{v}_i$  by solving  $1 + M$  (nonlinear) state problems and  $1 + 2N + 2M$  linear problems,  $i = 1, \dots, M, j = 1, \dots, N$ . For the eigenpairs,  $4(N + p)$  linear problems needs to be solved. By setting variation with respect to  $\lambda_j, \hat{v}_j$ , and  $\hat{u}_j$  to zero, we obtain similar expressions (50), (52), and (54) for the adjoint variables  $\psi_j^*, \hat{u}_j^*$ , and  $\hat{v}_j^*$ , where the constant  $\frac{1+2\beta\lambda_j}{2}$  is replaced by  $\frac{1+2\beta\lambda_j}{2} - \beta c$  where  $c$  is given by

$$c = \frac{1}{M} \sum_{i=1}^M Q(m_i) - Q(\bar{m}) - \langle m_i - \bar{m}, Q_m(\bar{m}) \rangle - \frac{1}{2} \langle m_i - \bar{m}, Q_{mm}(\bar{m})(m_i - \bar{m}) \rangle. \tag{63}$$

By setting variation with respect to  $\hat{v}_i, \hat{u}_i, u_i, v, u$  to vanish, we get the adjoints  $\hat{u}_i^*, \hat{v}_i^*, v_i, u^*, v^*$  by solving  $2 + 3M$  linear problems, which are defined similarly as that for the quadratic approximation, but with the additional correction terms in

**Table 1**

The computational cost for the evaluation of the cost functional  $\mathcal{J}$  and its  $z$ -gradient  $D_z \mathcal{J}$  in terms of the number of PDE solves for the five cases of approximation of  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$ .  $N$  is the number of samples/bases for the trace estimator;  $p$  is the oversampling factor ( $p = 5 \sim 10$ ) in Algorithm 2;  $M$  is the number of samples used in the sample average approximation or the Monte Carlo correction.

| Computational cost             | MC  | Linear | Quadratic     | Linear+MC | Quadratic+MC       |
|--------------------------------|-----|--------|---------------|-----------|--------------------|
| $\mathcal{J}$ (state PDE)      | $M$ | 1      | 1             | $1 + M$   | $1 + M$            |
| $\mathcal{J}$ (linear PDE)     | 0   | 1      | $1 + 4N + 4p$ | 1         | $1 + 4N + 4p + 2M$ |
| $D_z \mathcal{J}$ (linear PDE) | $M$ | 2      | $2 + 2N$      | $2 + M$   | $2 + 2N + 3M$      |

both the cost functional and the constraints, thus in total  $1 + M$  (nonlinear) state problems and  $3 + 6N + 4p + 5M$  linear problems to evaluate the  $z$ -gradient

$$\langle \tilde{z}, D_z \mathcal{J}_{\text{quad}}^{\text{MC}}(z) \rangle = \langle \tilde{z}, \partial_z \mathcal{L}_{\text{quad}}^{\text{MC}}(\text{states, adjoints}, z) \rangle, \tag{64}$$

where states and adjoints represent the state and adjoints for short.

#### 4.4. $z$ -gradient-based optimization method

To solve the PDE-constrained optimal control problem we use a gradient-based optimization method, since it only requires the ability to evaluate the cost functional (7) and its  $z$ -gradient. In particular, we use a BFGS method [62].

The five approximations described in this section all differ among each other in terms of accuracy and computational cost. Table 1 summarizes the cost, measured in number of (nonlinear) state and linear PDE solves, to evaluate cost functional and its  $z$ -gradient for each approximation. The linear approximation in (41) is the cheapest to compute but also the most crude, while the MC-corrected approximations in (58) and (61) are the most accurate but also the most expensive. The quadratic approximation (46) is, in many applications, more accurate than the linear approximation and, for highly nonlinear problems, is much cheaper to compute than the MC-corrected approximation. In fact, all the linear PDE solves in the trace approximation share the same left hand side  $\partial_{vu} \bar{r}$  (or its adjoint  $\partial_{uv} \bar{r}$ ), and therefore one can amortize the cost of computing an expensive preconditioner or sparse factorization over several solutions of the linearized PDE. It is worth noting that these different approximations can be exploited in a multifidelity framework [63] where the optimal solution of a cheaper approximation is used as initial guess for a more accurate (and more expensive) approximation. For example, in the numerical results section the optimal solution using the MC-corrected formulation are obtained by first solving the linear approximation in (41), then solving the quadratic approximation in (46) and finally (58) or (61). We remark that, in general, the approximations (41), (46), (58), (61) may be non-convex with respect to the control, even when the cost functional (7) is convex. In such cases, continuation from simple (e.g., deterministic or linear) approximations to more complex (e.g., quadratic with MC correction) ones are also helpful in finding a global, or at least a good local, optimum.

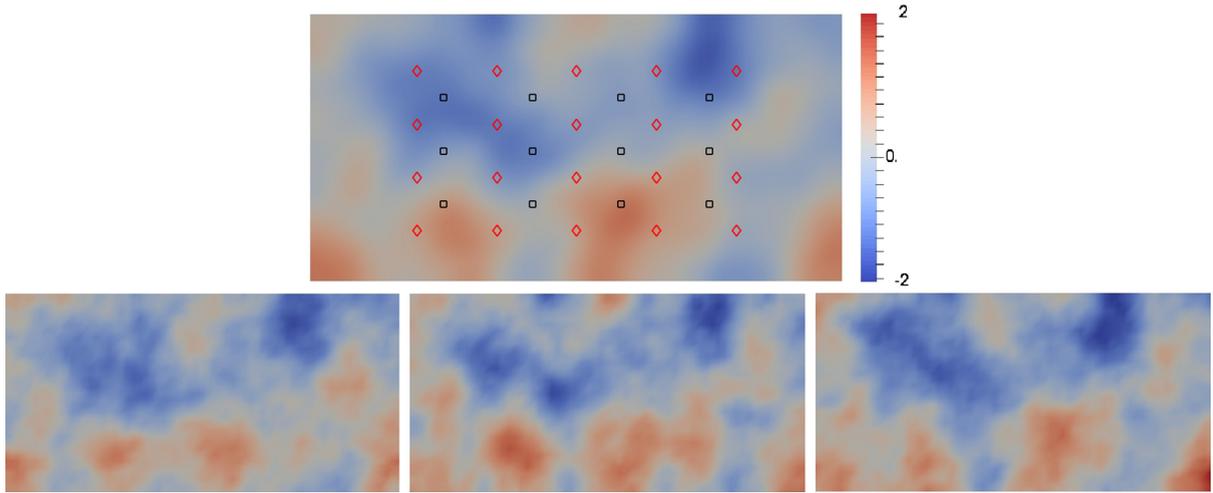
### 5. Numerical experiments

We consider two examples of PDE-constrained optimal control under uncertainty in which the uncertain parameter is the discretization of an infinite-dimensional random field. These examples aim to demonstrate the scalability, computational efficiency and accuracy of the approximation based on the Taylor approximation and their MC-corrected counterparts. In the first example we consider a fluid flow in a porous medium; the control is the fluid injection rate into the subsurface at specific well locations, the uncertain parameter is the permeability field (modeled as a log-normal random field), and the objective is to drive the pressures measured at the production wells to given target values. In the second example we seek an optimal velocity profile on the inlet boundary of a turbulent jet flow such that the jet width in a given cross-section is maximized; the governing state equations are the Reynolds-averaged Navier–Stokes equations (RANS) with an algebraic closure model and an ad-hoc nonlinear advection–diffusion equation with uncertain coefficients is introduced to model the intermittency of turbulence at the edges of the jet.

The numerical results presented in this section are obtained using `hippylib` [76], an extensible software framework for large scale PDE-based deterministic and Bayesian inverse problems. `hippylib` builds on `FEniCS` [57] (a parallel finite element library) for the discretization of the PDE and `PETSc` [7] for scalable and efficient linear algebra operations and solvers. Finally, we use the implementation of the L-BFGS algorithm in the Python library `SciPy`, specifically the `fmin_l_bfgs_b` routine (BFGS with limited memory and box constraint), to solve the optimization problem.

#### 5.1. Optimal source control of fluid flow in a porous medium

We consider an optimal source control problem—originally presented in [3]—motivated by flow in a subsurface porous medium, where the state problem is a linear elliptic equation describing single phase fluid flow in a porous medium:



**Fig. 1.** Top: mean  $\bar{m}$  of the Gaussian uncertain parameter  $m \sim \mathcal{N}(\bar{m}, \mathcal{C})$ . The red diamonds represent the locations of the injection wells, and the black squares represent the locations of the production wells. Bottom: random samples drawn from the Gaussian measure  $\mathcal{N}(\bar{m}, \mathcal{C})$ . (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$\begin{aligned} -\nabla \cdot (e^m \nabla u) &= Fz && \text{in } D, \\ u &= g && \text{on } \Gamma_D, \\ e^m \nabla u \cdot \mathbf{n} &= 0 && \text{on } \Gamma_N, \end{aligned} \tag{65}$$

where  $D$  is an open and bounded physical domain  $D \subset \mathbb{R}^d$  ( $d = 2, 3$ ) with Lipschitz boundary  $\partial D$ , and  $\Gamma_D, \Gamma_N \subset \partial D$  ( $\Gamma_D \cup \Gamma_N = \partial D$  and  $\Gamma_D \cap \Gamma_N = \emptyset$ ) denote respectively the Dirichlet and Neumann portion of  $\partial D$ .  $u$  is the state representing the pressure.  $m$  is an uncertain parameter representing the logarithm of the permeability field, which is assumed to be Gaussian  $\mathcal{N}(\bar{m}, \mathcal{C})$  with mean  $\bar{m}$  and covariance  $\mathcal{C}$ .  $F$  is a map from the control  $z$  to the source term  $Fz$  defined as

$$Fz = \sum_{i=1}^{n_c} z_i f_i, \tag{66}$$

where  $f_i, i = 1, \dots, n_c$ , are the mollified Dirac functions located at the  $n_c$  injection wells, and  $z = (z_1, \dots, z_{n_c})^\top \in \mathbb{R}^{n_c}$  is a control that represents the fluid injection rate. The control  $z$  takes admissible value in  $\mathcal{Z} = [z_{\min}, z_{\max}]^{n_c}$ , where  $z_{\min}$  and  $z_{\max}$  denote the minimum and maximum injection rate.

Let  $R_g$  denote the lifting of the boundary data  $g$  such that  $R_g \in H^1(D)$  and  $R_g|_{\Gamma_D} = g$ . The weak formulation of problem (65) reads: Given  $m \in \mathcal{M} = L^2(D)$  and  $z \in \mathcal{Z}$ , find  $\hat{u} = u - R_g \in \mathcal{U} := \{v \in H^1(D) : v|_{\Gamma_D} = 0\}$  such that

$$r(\hat{u}, v, m, z) = 0, \quad \forall v \in \mathcal{V}, \tag{67}$$

where  $\mathcal{V} = \mathcal{U}$ , and

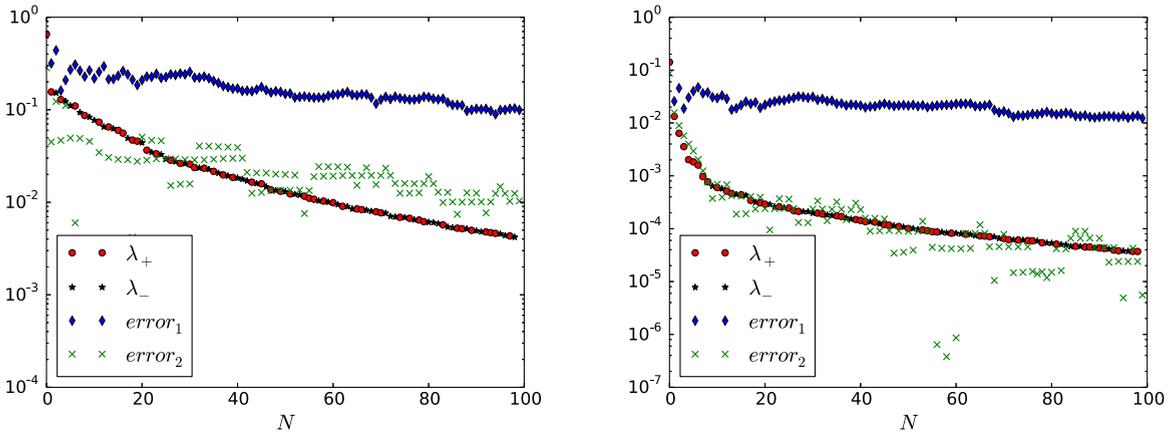
$$r(\hat{u}, v, m, z) = \int_D e^m \nabla(\hat{u} + R_g) \cdot \nabla v dx - \int_D Fz v dx. \tag{68}$$

The control objective measures the difference between the pressure (computed by solving the state problem (67)) and a target pressure distribution  $\bar{u} = (\bar{u}_1, \dots, \bar{u}_{n_p})$  at the production wells  $x^1, \dots, x^{n_p} \in D$ . Specifically,

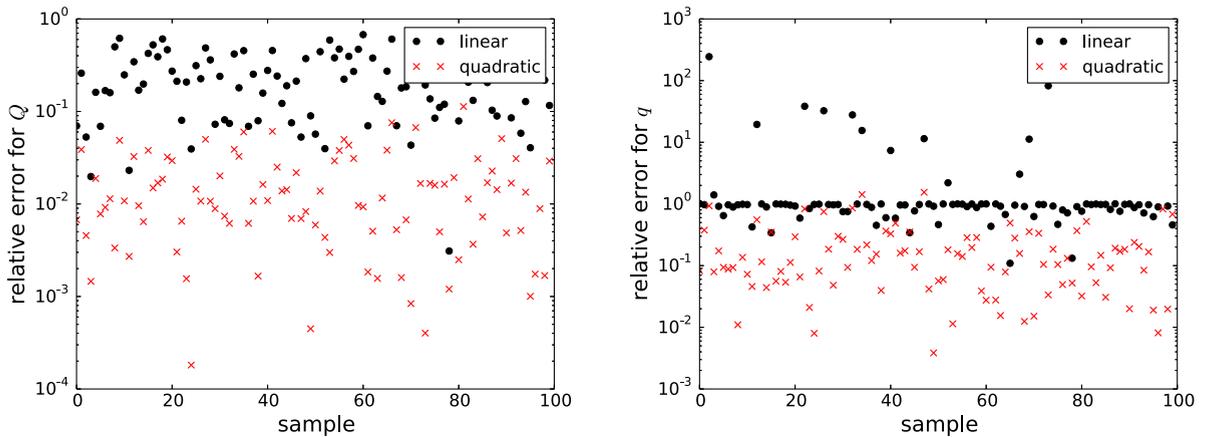
$$Q(\hat{u}) = \|\mathcal{B}\hat{u} - \bar{u}\|_2^2, \tag{69}$$

where  $\mathcal{B} : \mathcal{U} \rightarrow \mathbb{R}^{n_p}$  is a pointwise observation operator at the locations  $x^1, \dots, x^{n_p}$ .

Here, we set  $D = (0, 2) \times (0, 1)$ ,  $\Gamma_D = \{0, 2\} \times [0, 1]$ , and  $\Gamma_N = (0, 2) \times \{0, 1\}$ . We impose the Dirichlet data  $g = 2 - x_1$ . The mean field  $\bar{m}$  is shown in the top part of Fig. 1, in which locations for the 20 injection wells and the 12 production wells are also shown. For the covariance  $\mathcal{C}$  defined in (3), we specify  $\Theta = [1, 0; 0, 1]$ ,  $\alpha = 2$ ,  $\alpha_1 = 0.1$ , and  $\alpha_2 = 20$ . Three random samples of the Gaussian measure  $\mathcal{N}(\bar{m}, \mathcal{C})$  are shown in the bottom part of Fig. 1. In the cost functional (7), we set  $\beta = 1$  and use the control cost  $\mathcal{P}(z) = 10^{-5} \times \|z\|_2^2$ , where the parameters are chosen as in [3]. We use a finite element discretization with a structured uniform triangular mesh with  $65 \times 33$  vertices and piecewise linear elements for the discrete approximation of the state  $u$ , the adjoint  $v$ , and the uncertain parameter  $m$ , which results in 2145 dimensions for  $m$ . We specify the desired pressure  $\bar{u}_i = 3 - 8(x_1^i - 1)^2 - 4(x_2^i - 0.5)^2$  at the production wells  $x^i = (x_1^i, x_2^i)$ ,  $i = 1, \dots, 12$ , which are shown in the top part of Fig. 1. Moreover, we set the lower and upper bounds  $z_{\min} = 0$  and  $z_{\max} = 32$  for the control, i.e.,  $z \in [0, 32]^{20}$ .



**Fig. 2.** The decay of the generalized eigenvalues,  $\lambda_+$  for positive eigenvalues and  $\lambda_-$  for negative eigenvalues, and the errors, denoted as  $error_1$  and  $error_2$  for the trace estimators  $\widehat{T}_1$  and  $\widehat{T}_2$  with respect to the number  $N$  in (26) and (27). Left: at the initial control  $z = z_0$ ; right: at the optimal control  $z = z_{quad}^{MC}$  obtained by quadratic approximation with Monte Carlo correction.



**Fig. 3.** Relative errors at 100 random samples of the uncertain parameter for the linear and quadratic approximations of  $Q$  used in the computation of the mean  $\mathbb{E}[Q]$  (left) and of  $q = (Q - \bar{Q})^2$  for the variance  $\text{Var}[Q]$  (right). Here, the errors are show at the optimal control  $z = z_{quad}^{MC}$  obtained by quadratic approximation with Monte Carlo correction.

We consider the five different approximations of the cost functional presented in Sec. 3, and use the limited memory BFGS method with box constraints to solve the optimization problem, where the  $z$ -gradient of the cost functional is computed as in Sec. 4. We chose  $z_0 = (16, \dots, 16)$  as an initial guess for the control. We used  $M = 100$  random samples from  $\mathcal{N}(\bar{m}, \mathcal{C})$  to estimate the mean and variance of the control objective in both the sample average approximation presented in Sec. 2.4 and the variance reduction in Sec. 3.4.

For the computation of the trace appearing in the quadratic approximations (16), (33), and (35), we use the trace estimator  $\widehat{T}_2(\cdot)$  (27) by solving the generalized eigenvalue problem (28) using Algorithm 2, with  $k = 100$  and  $p = 10$ . We also test the Gaussian trace estimator  $\widehat{T}_1(\cdot)$  in (26). In Fig. 2, we plot the errors, denoted as  $error_1$  and  $error_2$  for the two trace approximations, respectively, against the reference value  $\sum_{j=1}^k \lambda_j$  with  $k = 140$  at both the initial control  $z = z_0$  and the optimal control  $z = z_{quad}^{MC}$  obtained by the quadratic approximation with Monte Carlo correction. We can observe that in both cases, the error for  $\widehat{T}_2(\mathcal{H})$  decays much faster than that for  $\widehat{T}_1(\mathcal{H})$  as a result of fast decaying eigenvalues, especially at the optimal control where more than two orders of magnitude improvement in accuracy is observed for  $N = 100$ . We remark that the slow decay of the error for  $\widehat{T}_1(\cdot)$  may demand a very large  $N$  if relatively high accuracy of the trace estimator is required, which leads to a large number of PDE solves at each step of the optimization procedure, thus diminishing the computational advantage of using the Taylor approximation. In the following test, we use  $\widehat{T}_2(\cdot)$  with  $N = 100$ .

Fig. 3 shows the relative errors at 100 random samples of the uncertain parameter  $m$  for the linear and quadratic approximations of  $Q$  in (69) used for the computation of  $\mathbb{E}[Q]$ , as well as of  $q = (Q - \bar{Q})^2$  in (32) and (35) for the computation of  $\text{Var}[Q]$ . We note that, for both quantities, the quadratic approximation is more accurate than the linear approximation for most of the realizations.

**Table 2**

Variance reduction by the linear and quadratic approximations at the initial control  $z_0$ , and the optimal controls  $z_{lin}^{MC}$  and  $z_{quad}^{MC}$ . The mean-square-error (MSE) with different numbers of random samples for the evaluation of the mean  $\mathbb{E}[Q]$  are reported.

| control         | # samples | $\mathbb{E}^{MC}[Q]$ | MSE(Q)   | MSE(Q - $Q_{lin}$ ) | MSE(Q - $Q_{quad}$ ) |
|-----------------|-----------|----------------------|----------|---------------------|----------------------|
| $z_0$           | 10        | 2.42e+01             | 1.10e+00 | 7.94e-03            | 1.52e-04             |
|                 | 100       | 2.54e+01             | 2.54e-01 | 4.16e-03            | 7.40e-05             |
| $z_{lin}^{MC}$  | 10        | 3.01e-01             | 1.10e-03 | 1.31e-03            | 1.69e-05             |
|                 | 100       | 2.89e-01             | 9.68e-05 | 7.49e-05            | 9.00e-07             |
| $z_{quad}^{MC}$ | 10        | 3.16e-01             | 1.11e-03 | 8.02e-04            | 1.05e-05             |
|                 | 100       | 2.98e-01             | 1.10e-04 | 9.98e-05            | 1.12e-06             |

**Table 3**

Variance reduction by the linear and quadratic approximations at the initial control  $z_0$ , and the optimal controls  $z_{lin}^{MC}$  and  $z_{quad}^{MC}$ . The MSE for the evaluation of the variance  $\text{Var}[Q]$ , where  $q = (Q - \bar{Q})^2$ ,  $q_{lin} = (Q_{lin} - \bar{Q})^2$ ,  $q_{quad} = (Q_{quad} - \bar{Q})^2$  are reported.

| control         | # samples | $\mathbb{E}^{MC}[q]$ | MSE(q)   | MSE(q - $q_{lin}$ ) | MSE(q - $q_{quad}$ ) |
|-----------------|-----------|----------------------|----------|---------------------|----------------------|
| $z_0$           | 10        | 1.32e+01             | 1.71e+01 | 9.38e-01            | 6.05e-03             |
|                 | 100       | 2.54e+01             | 1.35e+01 | 2.61e+00            | 3.09e-02             |
| $z_{lin}^{MC}$  | 10        | 1.99e-02             | 1.00e-04 | 9.35e-05            | 3.92e-06             |
|                 | 100       | 1.65e-02             | 1.32e-05 | 1.22e-05            | 2.44e-07             |
| $z_{quad}^{MC}$ | 10        | 2.38e-02             | 1.82e-04 | 1.79e-04            | 8.07e-07             |
|                 | 100       | 1.99e-02             | 1.66e-05 | 1.65e-05            | 3.81e-07             |

Since the quadratic approximation is very close and strongly correlated to the true value, we may expect large variance reduction by using it as a control variate. In fact, the correlation coefficient between the quadratic approximation and the true value is very close to one, 0.996 for  $Q$  and 0.998 for  $q$  evaluated using the 100 samples. On the contrary, the linear approximation is poorly correlated to the true value (with correlation coefficient  $-0.211$  for  $Q$  and  $0.718$  for  $q$ ), thus leading to negligible variance reduction. In Tables 2 and 3, we report the effect of the variance reduction by the linear and quadratic approximations for the computation of the mean  $\mathbb{E}[Q]$  and variance  $\text{Var}[Q]$ , respectively. This reduction is measured by the mean-square-error (MSE), which is defined (e.g., for  $Q$ ) as

$$\text{MSE}(Q) = \frac{\text{Var}_p[Q]}{M}, \tag{70}$$

where  $\text{Var}_p[Q]$  is the population variance of  $Q$  at  $M$  random samples.

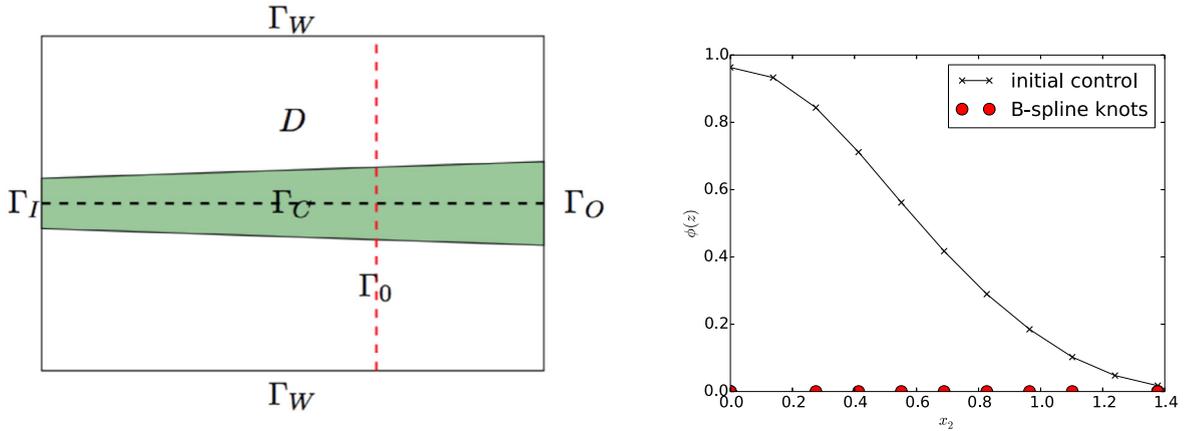
In all cases, we can see that the quadratic approximation achieves more significant variance reduction than the linear approximation. This means that the number of samples needed to achieve a target MSE is significantly smaller for the quadratic approximation than for the linear. For example, at the initial control  $z_0$ , the quadratic approximation results in a speed-up (defined as the ratio between the MSE of the Monte Carlo estimator and the MSE of the correction) of order 1000X and 100X in the estimation of  $\mathbb{E}[Q]$  and  $\text{Var}[Q]$ , respectively. On the other hand, the linear approximation provides a speed-up of order 100X and 10X respectively. In addition, at the optimal controls  $z_{lin}^{MC}$  and  $z_{quad}^{MC}$ , the quadratic approximation can still achieve a  $\sim 100X$  speed up for evaluation of both the mean and variance, while the linear approximation is too poor to achieve meaningful variance reduction. Note that the mean and variance are both much smaller at the optimal control than those at the initial control, confirming that the cost functional is significantly reduced by solving the minimization problem (9).

In this example, we see that the trace estimator using the generalized eigenvalues computed by Algorithm 2 is much more accurate than that using the Gaussian randomized (Monte Carlo) trace estimator thanks to the fast decay of the generalized eigenvalues. Moreover, using the Taylor approximation, in particular the quadratic approximation, as a control variate results in over two orders of magnitude in variance reduction for the evaluation of the cost functional.

### 5.2. Optimal boundary control of a turbulent jet flow

We consider a turbulent jet flow modeled by the Reynolds-averaged Navier–Stokes equations coupled with a nonlinear stochastic advection-diffusion equation that serves as a notional model of uncertainty in the turbulent viscosity  $\nu_t$ . Specifically, the governing equations read

$$\begin{aligned} -\nabla \cdot \left( (\nu + \nu_t) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) \right) + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p &= 0, & \text{in } D, \\ \nabla \cdot \mathbf{u} &= 0, & \text{in } D, \\ -\nabla \cdot \left( (\nu + (\gamma + e^m) \nu_{t,0}) \nabla \gamma \right) + \mathbf{u} \cdot \nabla \gamma - \frac{1}{2} \frac{\mathbf{u} \cdot \mathbf{e}_1}{x_1 + b} \gamma &= 0, & \text{in } D, \end{aligned} \tag{71}$$



**Fig. 4.** Left: sketch of the physical domain of the turbulent jet flow, with inlet boundary  $\Gamma_I$ , outlet boundary  $\Gamma_O$ , top and bottom wall  $\Gamma_W$ , center axis  $\Gamma_C$ , and the cross-section  $\Gamma_0$ . The computational domain  $D$  is the symmetric top half of the physical domain. Right: B-spline interpolation of the inlet velocity profile used as an initial guess, and the distribution of the B-spline knots.

with the boundary conditions

$$\begin{aligned}
 \sigma_n(\mathbf{u}) \cdot \boldsymbol{\tau} &= 0, \quad \mathbf{u} \cdot \mathbf{n} + \chi_W \phi(z) = 0, & \text{on } \Gamma_I, \\
 \sigma_n(\mathbf{u}) \cdot \mathbf{n} &= 0, \quad \mathbf{u} \cdot \boldsymbol{\tau} = 0, & \text{on } \Gamma_O \cup \Gamma_W, \\
 \sigma_n(\mathbf{u}) \cdot \boldsymbol{\tau} &= 0, \quad \mathbf{u} \cdot \mathbf{n} = 0, & \text{on } \Gamma_C, \\
 \gamma - \gamma_0 &= 0, & \text{on } \Gamma_I \cup \Gamma_W, \\
 \sigma_n^\gamma(\gamma) \cdot \mathbf{n} &= 0, & \text{on } \Gamma_O \cup \Gamma_C.
 \end{aligned} \tag{72}$$

Here,  $D = (0, 30) \times (0, 10)$  is the computational domain (the symmetric upper half of the physical domain for the jet flow) as shown in the left part of Fig. 4.  $\mathbf{e}_1 = (1, 0)$ ,  $\boldsymbol{\tau}$  and  $\mathbf{n}$  are the unit tangential vector and unit normal vector pointing outside the domain  $D$  along the boundary  $\partial D$ ,  $\sigma_n(\mathbf{u}) = (\nu + \nu_t)(\nabla \mathbf{u} + \nabla \mathbf{u}^\top) \cdot \mathbf{n}$  is the traction vector, and  $\sigma_n^\gamma(\gamma) = (\nu + (\gamma + e^m)\nu_{t,0})\nabla \gamma \cdot \mathbf{n}$  is the normal flux of  $\gamma$ . The control portion of the inflow boundary is denoted by the indicator function  $\chi_W$  with  $\chi_W(x) = 1$  if  $x_2 \in [0, W]$  and  $\chi_W(x) = 0$  if  $x_2 \in (W, 10)$ . Here we take  $W = 1.5$ . The state variables  $\mathbf{u} = (u_1, u_2)$  and  $p$  are the time-averaged velocity and pressure, and  $\nu_t = \gamma \nu_{t,0}$  is the turbulent viscosity, where  $\nu_{t,0}$  represents an algebraic closure model and  $\gamma$  is the indicator function that is close to one near the centerline  $\Gamma_C$  and zero outside the jet region. Based on a dimensional analysis of the turbulent planar jet, the turbulent viscosity along the jet centerline is given by

$$\nu_{t,0} = C\sqrt{M}(x_1 + aW)^{1/2}, \tag{73}$$

where  $M = \int_{\Gamma_I} \|\mathbf{u}\|^2 ds$  and  $W$  are momentum and jet width at the inlet boundary,  $C$  and  $a$  are two adimensional parameters that need to be calibrated against either experimental or direct Navier-Stokes simulation (DNS) data. The indicator function  $\gamma$  satisfies the nonlinear stochastic advection-diffusion equation in (71) with the boundary data

$$\gamma_0 = 0.5 - 0.5 \tanh\left(5\left(\frac{30 - x_1}{30}\right)(x_2 - 1 - 0.5x_1)\right). \tag{74}$$

The uncertain parameter  $m$  is a Gaussian random field with distribution  $\mathcal{N}(\bar{m}, \mathcal{C})$ , for which we set the mean  $\bar{m}$  as a constant field and the covariance as in (3) with  $\alpha = 2$ ,  $\alpha_1 = \alpha_2 = 0.5$ , and  $\Theta = [15, 0; 0, 1]$ . Two samples drawn from  $\mathcal{N}(0, \mathcal{C})$  are shown in Fig. 5. In general, the parameters  $(C, a, \bar{m}, b)$  are also random variables whose distributions can be inferred from experimental or DNS data. Here, for simplicity, we consider only the uncertainty in  $m$  and take deterministic values for  $(C, a, \bar{m}, b) = (0.012, 5.29, -0.47, 9.58)$ .<sup>1</sup>

The function  $\phi(z)$  in (72) is the horizontal component of the velocity profile on the inlet boundary  $\Gamma_I$ , which is represented by B-splines as (we use  $t$  for  $x_2$  here)

$$\phi(z)(t) = \sum_{k=1}^K z_k B_{k,n}(t), \tag{75}$$

<sup>1</sup> These values are the maximum a posteriori estimates obtained by solving a Bayesian calibration problem where the misfit functional is the  $L^2$  distance between the velocity of (71) and the DNS data provided in [48].

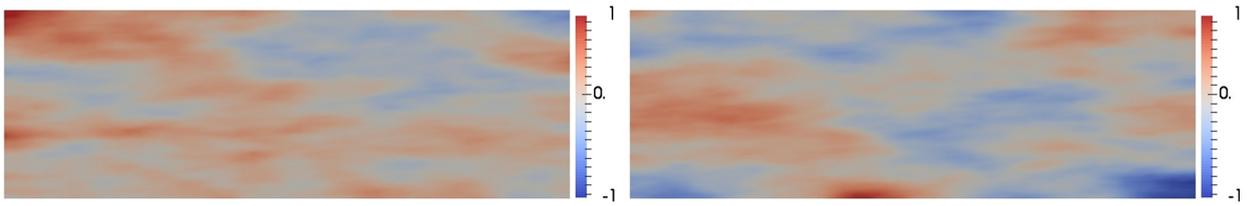


Fig. 5. Two random samples drawn from the Gaussian distribution  $\mathcal{N}(0, C)$  with  $C$  given in (3), where  $\alpha = 2$ ,  $\alpha_1 = \alpha_2 = 0.5$ ,  $\Theta = [15, 0; 0, 1]$ .

where the coefficient vector  $z := (z_1, \dots, z_K) \in [0, 1]^K$  is the control and  $B_{k,n}(t)$  are B-spline basis functions recursively defined as

$$B_{k,n}(t) = \frac{t - t_k}{t_{k+n} - t_k} B_{k,n-1}(t) + \frac{t_{k+n+1} - t}{t_{k+n+1} - t_{k+1}} B_{k+1,n-1}(t), \tag{76}$$

where

$$B_{k,0}(t) = \begin{cases} 1 & \text{if } t_k \leq t < t_{k+1}; \\ 0 & \text{otherwise.} \end{cases} \tag{77}$$

Here we take  $K = 11$  and use the B-spline of order  $n = 3$  with 15 knots (4 repetitions at each endpoint), as depicted on the right of Fig. 4. A B-spline representation of the inflow velocity profile used for the initial control is also shown.

The control objective is the jet width  $Q$  along the cross-section  $\Gamma_0$  at  $x_1 = 20$  shown in Fig. 4, and is defined as

$$Q(\mathbf{u}) = \frac{1}{\delta_1^0(\mathbf{u})} \int_{\Gamma_0} \mathbf{u} \cdot \mathbf{e}_1 dx_2, \tag{78}$$

where  $\delta_1^0(\mathbf{u}) = \mathbf{u}(x^0) \cdot \mathbf{e}_1$  with  $x^0 = (20, 0)$ . The objective is to maximize the expected value of the jet width  $Q$ , i.e., to minimize  $\mathbb{E}[-Q]$  subject to the state problem (71) and the constraint that the inlet momentum  $M$  is prescribed. Specifically, we consider the cost functional

$$\mathcal{J}(z) = \mathbb{E}[-Q] + \beta \text{Var}[Q] + \mathcal{P}(z). \tag{79}$$

Here we set  $\beta = 10^3$  to balance the mean and variance, and define the penalty term  $\mathcal{P}(z)$  as

$$\mathcal{P}(z) = \beta_1 \left( \int_{\Gamma_I} (\phi(z))^2 dx - M \right)^2 + \beta_2 \int_{\Gamma_I} |\nabla \phi(z)|^2 dx, \tag{80}$$

where the first term penalizes violations of the inlet momentum constraint (with a relatively large  $\beta_1 = 10^3$ ), and the second term controls the regularity of the inlet velocity ( $\beta_2 = 1$ ).

The weak formulation of the turbulence model is given by: Given the realization  $m \in \mathcal{M} = L^2(D)$  and control  $z \in \mathcal{Z} = \mathbb{R}^K$ , find  $u \in \mathcal{U}$  such that

$$r(u, v, m, z) = 0, \quad \forall v \in \mathcal{V}, \tag{81}$$

where we denote the state with  $u = (\hat{\mathbf{u}}, p, \hat{\gamma}) \in \mathcal{U}$  and adjoint with  $v = (\mathbf{v}, q, \eta) \in \mathcal{V}$ . The space for the state and adjoint variables are  $\mathcal{U} = \mathcal{V} = U \times P \times V$ , where  $U = \{\mathbf{v} \in (H^1(D))^2 : \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \Gamma_C \text{ and } \mathbf{v} \cdot \boldsymbol{\tau} = 0 \text{ on } \Gamma_O \cup \Gamma_W\}$ ,  $P = L^2(D)$ , and  $V = \{\gamma \in H^1(D) : \gamma|_{\Gamma_W \cup \Gamma_I} = 0\}$ . Let  $R_0$  be the lifting of the Dirichlet boundary data for velocity on  $\Gamma_C \cup \Gamma_O \cup \Gamma_W$ , we have  $\mathbf{u} = \hat{\mathbf{u}} + R_0$ . Let  $R_{\gamma_0}$  be the lifting of the boundary data  $\gamma_0$  on  $\Gamma_I \cup \Gamma_W$ , we have  $\gamma = \hat{\gamma} + R_{\gamma_0}$ .

Specifically, the weak form in (81) consists of three terms and reads

$$r(u, v, m, z) = \text{Model}(u, v, m) + \text{Stabilization}(u, v) + \text{Nitsche}(u, v, m, z). \tag{82}$$

The first term represents the PDE model in weak form given by

$$\begin{aligned} \text{Model}(u, v, m) = & \int_D (v + v_t) 2\mathbb{S}(\mathbf{u}) \cdot \mathbb{S}(\mathbf{v}) dx + \int_D [(\mathbf{u} \cdot \nabla) \mathbf{u}] \cdot \mathbf{v} dx - \int_D p \nabla \cdot \mathbf{v} dx + \int_D q \nabla \cdot \mathbf{u} dx \\ & + \int_D (v + (\gamma + e^m) v_{t,0}) \nabla \gamma \cdot \nabla \eta dx + \int_D [\mathbf{u} \cdot \nabla \gamma] \eta dx - \int_D \frac{1}{2} \frac{\mathbf{u} \cdot \mathbf{e}_1}{x_1 + b} \gamma \eta dx, \end{aligned} \tag{83}$$

where  $\mathbb{S}(\mathbf{u}) = (\nabla \mathbf{u} + \nabla \mathbf{u}^\top)/2$  denotes strain tensor.

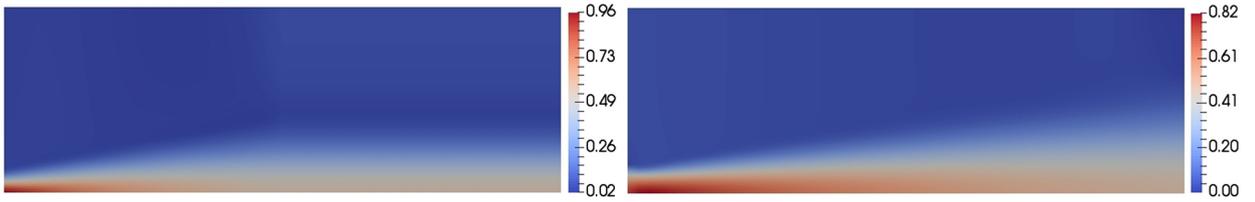


Fig. 6. The velocity field of the turbulent jet flow obtained solving the state model for the initial (left) and optimal (right) inlet velocity profile. Specifically, the optimal profile was obtained using the quadratic approximation with Monte Carlo correction.

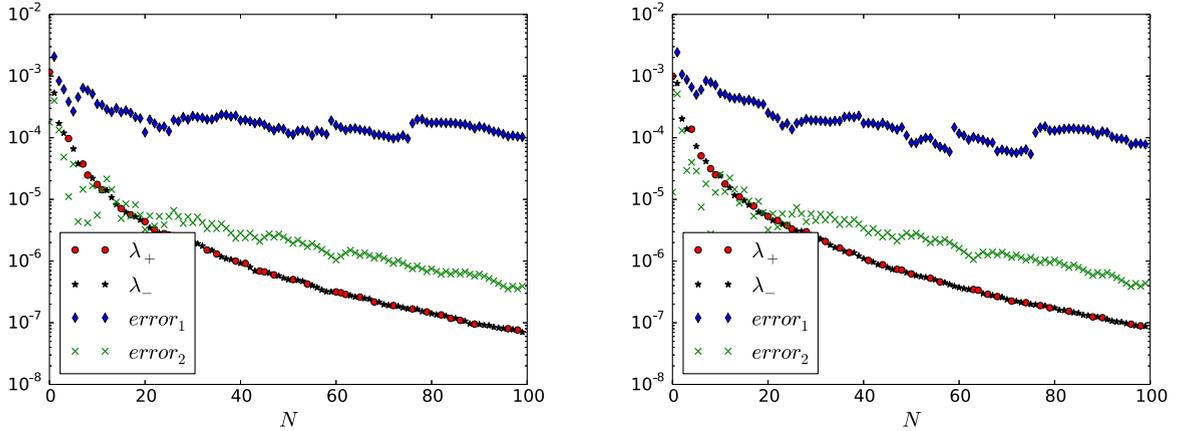


Fig. 7. The decay of the generalized eigenvalues,  $\lambda_+$  for positive eigenvalues and  $\lambda_-$  for negative eigenvalues, and the errors, denoted as  $error_1$  and  $error_2$  for the trace estimators  $\hat{T}_1$  and  $\hat{T}_2$  with respect to the number  $N$  in (26) and (27). Left: the initial control  $z = z_0$ ; right: the optimal control  $z = z_{quad}^{MC}$  obtained by quadratic approximation with Monte Carlo correction.

The second term of (81) represents stabilization by the Galerkin Least-Squares (GLS) for the momentum and mass conservation equations, and by streamline diffusion for the nonlinear advection diffusion equation, i.e.,

$$\text{Stabilization}(u, v) = \int_D \tau_1 L_1(u) \cdot D_u L_1(u)(v) dx + \int_D \tau_2 (\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{v}) dx + \int_D \tau_3 (\mathbf{u} \cdot \nabla \gamma)(\mathbf{u} \cdot \nabla \eta) dx, \quad (84)$$

where  $L_1(u)$  represents the strong form of the residual of the momentum equation (line 1) of (71), and  $\tau_1$ ,  $\tau_2$  and  $\tau_3$  are properly chosen stabilization parameters associated with the local Péclet number.

The third term of (81) represents the weak imposition of the inlet velocity profile (i.e., the control) by Nitsche’s method [9] and reads

$$\begin{aligned} \text{Nitsche}(u, v, m, z) = & C_d \int_{\Gamma_I} h^{-1} (v + v_t)(\mathbf{u} \cdot \mathbf{n} + \chi_W \phi(z))(\mathbf{v} \cdot \mathbf{n}) ds \\ & - \int_{\Gamma_I} (\sigma_n(\mathbf{u}) \cdot \mathbf{n})(\mathbf{v} \cdot \mathbf{n}) + (\sigma_n(\mathbf{v}) \cdot \mathbf{n})(\mathbf{u} \cdot \mathbf{n} + \chi_W \phi(z)) ds, \end{aligned} \quad (85)$$

where the first term guarantees the coercivity of the form and the second term arises from integration by parts. The constant  $C_d$  is taken as  $C_d = 10^5$ , and  $h$  is a local element edge length along the Dirichlet boundaries. This weak imposition of the inlet velocity is necessary to reveal the boundary control in the optimization formulation.

We discretize the problem using finite element with piecewise polynomials (P2, P1, P1) for the state  $u$  and P1 for the uncertain parameter  $m$ . We triangulate the computational domain using an anisotropic mesh, locally refined in the jet region, with 14,400 triangular elements (60 elements along the  $x_1$ -direction and 120 along the  $x_2$ -direction). After discretization, we obtain an optimization problem in which the dimension of the control, state, and uncertain parameter are 11,73084, and 7381, respectively. As in the previous example, we use a BFGS method with bound constraints to solve the optimization problem (9) with the cost functional defined in (79) and the constraint defined by the state problem (81). At each BFGS iteration, we solve the fully-coupled nonlinear state problem using Newton’s method with an LU factorization of the Jacobian operator. To compute all the derivatives involved in the  $z$ -gradient presented in Sec. 4 we exploit the symbolic differentiation capabilities of FEniCS [57]. Fig. 6 shows that the turbulent jet is more spread out with larger jet width at the optimal value of the control compared to that at the initial control.

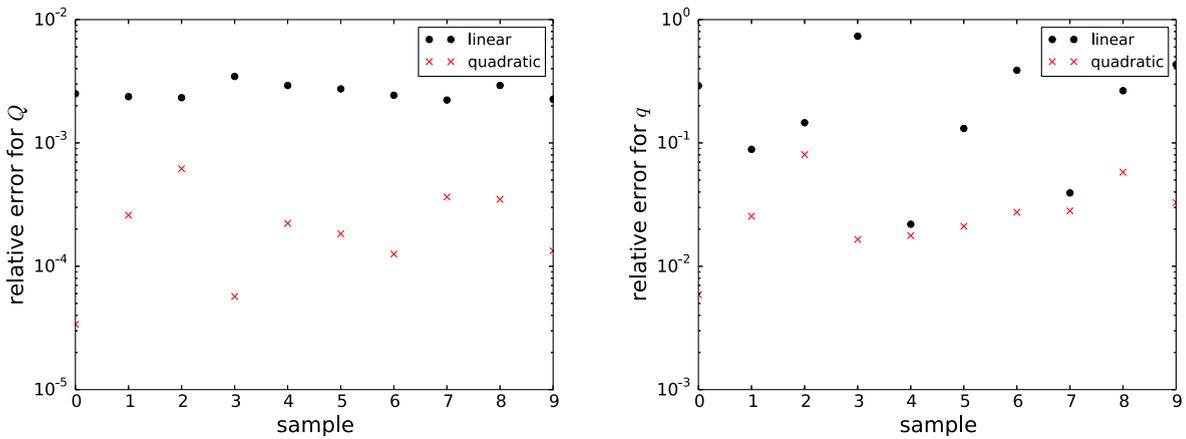


Fig. 8. Relative errors at 10 random samples of the linear and quadratic approximations of  $Q$  (for the computation of the mean  $\mathbb{E}[Q]$  (left)) and of  $q = (Q - \bar{Q})^2$  (for the variance  $\text{Var}[Q]$  (right)). Here, the errors are shown at the optimal control  $z = z_{\text{quad}}^{\text{MC}}$ .

Table 4

Variance reduction by the linear and quadratic approximations at the optimal control  $z_{\text{lin}}$  for different uncertain parameter dimensions. The mean-square-error (MSE) with 10 random samples for the evaluation of the mean  $\mathbb{E}[Q]$  are reported.

| parameter dimension | $\mathbb{E}^{\text{MC}}[-Q]$ | $\text{MSE}(Q)$ | $\text{MSE}(Q - Q_{\text{lin}})$ | $\text{MSE}(Q - Q_{\text{quad}})$ |
|---------------------|------------------------------|-----------------|----------------------------------|-----------------------------------|
| 1,891               | $-1.71e+00$                  | $7.40e-06$      | $2.68e-08$                       | $1.81e-09$                        |
| 7,381               | $-1.59e+00$                  | $7.94e-06$      | $1.57e-07$                       | $1.46e-08$                        |
| 29,161              | $-1.44e+00$                  | $3.82e-06$      | $7.23e-08$                       | $1.66e-08$                        |
| 115,921             | $-1.42e+00$                  | $9.47e-06$      | $6.91e-08$                       | $3.06e-08$                        |
| 462,241             | $-1.41e+00$                  | $8.85e-06$      | $9.00e-08$                       | $1.78e-08$                        |

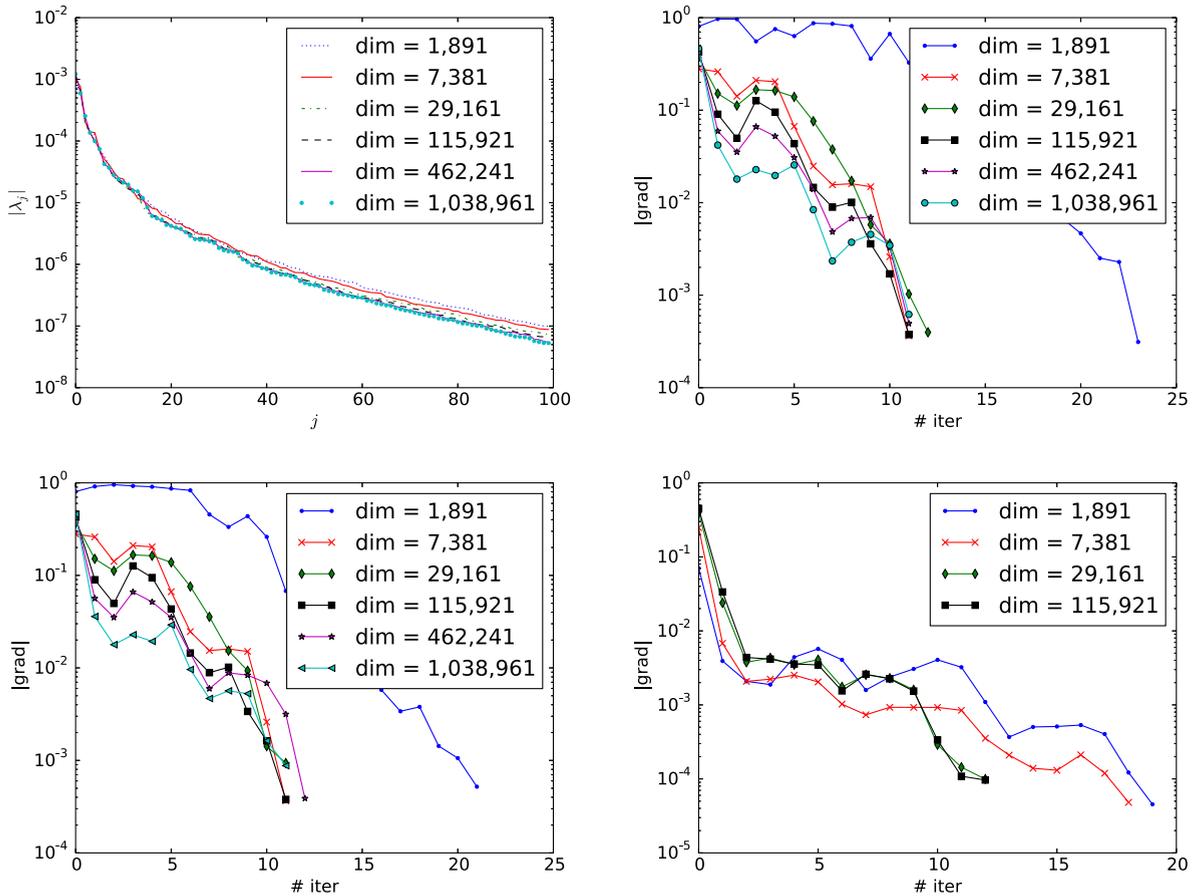
Table 5

Variance reduction by the linear and quadratic approximations at the optimal control  $z_{\text{lin}}$  for different uncertain parameter dimensions. The MSE with test 10 random samples for the evaluation of the variance are reported, where  $q = (Q - \bar{Q})^2$ ,  $q_{\text{lin}} = (Q_{\text{lin}} - \bar{Q})^2$ ,  $q_{\text{quad}} = (Q_{\text{quad}} - \bar{Q})^2$ .

| parameter dimension | $\mathbb{E}^{\text{MC}}[q]$ | $\text{MSE}(q)$ | $\text{MSE}(q - q_{\text{lin}})$ | $\text{MSE}(q - q_{\text{quad}})$ |
|---------------------|-----------------------------|-----------------|----------------------------------|-----------------------------------|
| 1,891               | $8.05e-05$                  | $9.37e-10$      | $1.76e-11$                       | $8.77e-13$                        |
| 7,381               | $8.13e-05$                  | $1.15e-09$      | $8.87e-12$                       | $1.48e-12$                        |
| 29,161              | $5.60e-05$                  | $6.59e-10$      | $3.39e-11$                       | $4.04e-12$                        |
| 115,921             | $1.07e-04$                  | $1.82e-09$      | $7.04e-11$                       | $2.08e-11$                        |
| 462,241             | $8.96e-05$                  | $1.37e-09$      | $7.78e-12$                       | $1.33e-11$                        |

For the computation of the trace in the quadratic approximation we use both the Gaussian trace estimator  $\widehat{T}_1(\cdot)$  in (26), and the trace estimator  $\widehat{T}_2(\cdot)$  in (27) obtained by solving the generalized eigenvalue problem (28). Fig. 7 displays the decay of the eigenvalues as well as the approximation error of the two estimators for both the initial control  $z_0$  and the optimal control  $z_{\text{quad}}^{\text{MC}}$ . The rapid decay of the eigenvalues (four orders of magnitude reduction in the first 100 out of 7381 eigenvalues) of the covariance-preconditioned Hessian reflects the intrinsic low-dimensionality of the map from the uncertain parameter field  $m$  to the control objective  $Q$ . This is despite the complexity and nonlinearity of the state equations. The reference value for the trace  $\text{tr}(\mathcal{H})$  was computed (with a precision of  $10^{-8}$ ) as the sum of the first 140 dominant eigenvalues of  $\mathcal{H}$  using Algorithm 2 with  $k = 140$  and  $p = 10$ . It is evident that the second trace estimator achieves much faster decay of the error than the Gaussian trace estimator, with a gain in accuracy of about three orders of magnitude when using  $N = 100$  samples/eigenvalues. In the numerical test, we use the second trace estimator with  $N = 20$ , which already provides sufficiently accurate results. It is worth noticing that, in contrast, the Gaussian trace estimator would have required about  $N = 10^4$  samples to achieve the same accuracy.

Fig. 8 shows the relative difference between the true quantity of interest and the linear and quadratic approximations evaluated at 10 random samples for both  $Q$  and  $q = (Q - \bar{Q})^2$ . It is clear that the quadratic approximation is more accurate than the linear approximation for both quantities. Nevertheless both approximations lead to drastic reduction in the variance of the Monte Carlo correction with respect to plain Monte Carlo estimator, as demonstrated by the MSE reported in Tables 4 and 5. This variance reduction translates to greater than 100X speed up in the computation relative to the plain Monte Carlo estimator for the same accuracy (MSE).



**Fig. 9.** Top-left: decay of the eigenvalues of the covariance-preconditioned Hessian (i.e., the generalized eigenvalues (in absolute value) of (28)) with different uncertain parameter dimensions (dim) at the optimal control obtained by linear approximation. Decay of the gradient with respect to the number of BFGS iterations (# iter), with linear approximation (top-right), quadratic approximation (bottom-left), and linear approximation with Monte Carlo correction (bottom-right). Smaller dimensionality is shown for the last case due to constraints on the computing time.

To investigate the scalability of the proposed algorithm, we consider up to six levels of uniform mesh refinement:  $30 \times 60$ ,  $60 \times 120$ ,  $120 \times 240$ ,  $240 \times 480$ ,  $480 \times 960$ , and  $720 \times 1440$ . The dimension of the uncertain parameter  $m$  ranges from 1,891 on the coarsest mesh to 1,038,961 on the finest. For the linear and quadratic approximation, we solve the optimization problem using the BFGS algorithm with convergence defined by an absolute tolerance of  $10^{-3}$  for the  $\ell^\infty$  norm of the projected gradient. We also use a mesh continuation technique to define initial guesses for the control: on the coarsest mesh we use the velocity profile in Fig. 4 (right), and, on the finer meshes, we use the optimal control obtained from the previous (coarser) mesh. For the linear approximation with Monte Carlo correction, we set the BFGS absolute tolerance to  $10^{-4}$  and we use the solution of the linear approximation (at the same mesh resolution) as the initial guess.

The scalability of our approach with respect to the uncertain parameter dimension depends on three properties, namely the dimension-independent behavior of (i) the spectral of the covariance preconditioned Hessian, (ii) the convergence rate of BFGS, and (iii) the variance of estimators for the Monte Carlo correction of the objective function. The top-left plot of Fig. 9 demonstrates (i): the eigenvalues of the covariance-preconditioned Hessian of the control objective (i.e., the generalized eigenvalues of problem (28)) at the optimal control obtained by linear approximation exhibit the same fast decay independent of the uncertain parameter dimension (ranging from a thousand to a million), thus indicating that approximating the trace by a randomized eigensolver is scalable with respect to the uncertain parameter dimension. The other three plots of Fig. 9 provide numerical evidence of the mesh independent convergence of the BFGS algorithm, property (ii), for this particular problem. In the case of the linear and quadratic approximations using mesh continuation, the norm of the z-gradient drops below the prescribed tolerance within 12 iterations for all mesh resolutions except the coarsest one (which represents a pre-asymptotic result); in the case of the linear approximation with Monte Carlo correction, BFGS converges within 19 iterations for all mesh resolutions using the solution of the linear approximation as initial guess. This implies that the gradient-based BFGS algorithm is also scalable with respect to the uncertain parameter dimension. Finally, Tables 4 and 5 indicate that the variance of the Monte Carlo correction, i.e., property (iii), using Taylor approximation as control variates is independent of the uncertain parameter dimension, and, furthermore, it allows greater than 100X speed up with respect to plain Monte Carlo estimator to achieve the same accuracy (MSE). Combining the three dimension-independent properties,

we conclude that the Taylor approximation and variance reduction stochastic optimization method is overall scalable with respect to the uncertain parameter dimension.

## 6. Conclusions

In this work we proposed a scalable computational framework to solve PDE-constrained optimal control under uncertainty where the cost functional involves the mean and variance of the control objective and the state equation is governed by a (nonlinear) PDE with high-dimensional uncertain parameter. Specifically, we advocate the use of (low-order) Taylor expansions of the control objective with respect to the uncertain parameter either to directly approximate the moments of the control objective or as control variates to reduce the variance of Monte Carlo estimator. The mean and variance expressions of the (linear and quadratic) Taylor expansions involve the trace of the covariance-preconditioned Hessian of the control objective with respect to the uncertain parameter. Randomized algorithms for solution of generalized eigenvalue problems [69,70,76] allow for accurate and efficient approximation of this trace and only require computing the action of the covariance-preconditioned Hessian on a number of random directions that depends on its numerical rank. As we showed in the numerical results, this approach is more efficient and accurate than the Gaussian trace estimator when the eigenvalues of the covariance-preconditioned Hessian exhibit fast decay, which is true when the control objective is only sensitive to a limited number of directions in the uncertain parameter space. This is often true regardless of the complexity of the state problem. Moreover, when the possibly biased Taylor (linear and quadratic) approximation is not sufficiently accurate, we can use it as a control variate and correct the mean and variance by Monte Carlo estimator of the remainder of the Taylor expansion, which was shown to achieve considerable computational cost reduction (by several orders of magnitude) compared to a plain Monte Carlo estimator. We have demonstrated the scalability of our method with respect to the uncertain parameter dimension in three aspects: the decay of the eigenvalues of the covariance-preconditioned Hessian of the control objective, the variance of the remainder of the Taylor expansion, and the number of optimization iterations. All three properties do not depend on the nominal dimension of the uncertain parameter but only on the intrinsic dimension of the uncertain parameter, i.e. on the number of directions the control objective is sensitive to. Moreover, by using the randomized eigensolver for trace estimator, Taylor approximation as control variate for variance reduction, and quasi-Newton method with a Lagrangian formalism for the optimization, we showed that the complexity—measured in number of PDE solves—scales linearly with respect to the intrinsic dimensionality to achieve target (high) accuracy, several orders of magnitude more accurate than a plain Monte Carlo estimator. This is illustrated numerically by solving an optimal boundary control governed by a nonlinear PDE model for turbulent jet flow with infinite-dimensional uncertain parameter field that characterizes the turbulent viscosity. Hence, the proposed computational framework for PDE-constrained optimal control under uncertainty is demonstrated scalable, accurate and efficient.

Several extensions and further developments of the proposed computational framework are worth pursuing. First, we intend to investigate higher-order Taylor expansions beyond the quadratic approximation when the later becomes poor. Second, another research direction is to pursue more general risk measures than the mean-variance measure, such as VaR and conditional VaR [50,73]. Third, further computational savings can be achieved in combination with model reduction for the state PDE and the linearized PDEs [30,77]. Fourth, one can replace the Monte Carlo estimator in the variance reduction by a sparse quadrature [24] to possibly achieve faster convergence of the approximation error. Fifth, application of the Taylor approximation and variance reduction in a multifidelity framework [63] may lead to additional computational saving. Finally, one can apply the scalable computational framework to other challenging control and design problems, such as optimal design of metamaterials under uncertainty [27].

## Acknowledgement

This work was supported by DARPA contract W911NF-15-2-0121, NSF grants CBET-1508713 and ACI-1550593, and DOE grants DE-SC0010518 and DE-SC0009286. We thank Robert Moser, Todd Oliver, and Myoungkyu Lee for the construction of the notional turbulent jet flow optimal boundary control under uncertainty problem.

## References

- [1] Alen Alexanderian, Noemi Petra, Georg Stadler, Omar Ghattas, A-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems with regularized  $\ell_0$ -sparsification, *SIAM J. Sci. Comput.* 36 (5) (2014) A2122–A2148.
- [2] Alen Alexanderian, Noemi Petra, Georg Stadler, Omar Ghattas, A fast and scalable method for A-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems, *SIAM J. Sci. Comput.* 38 (1) (2016) A243–A272.
- [3] Alen Alexanderian, Noemi Petra, Georg Stadler, Omar Ghattas, Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations, *SIAM/ASA J. Uncertain. Quantificat.* 5 (1) (2017) 1166–1192, arXiv preprint: arXiv:1602.07592.
- [4] Ahmad Ahmad Ali, Elisabeth Ullmann, Michael Hinze, Multilevel Monte Carlo analysis for optimal control of elliptic PDEs with random coefficients, *SIAM/ASA J. Uncertain. Quantificat.* 5 (1) (2017) 466–492.
- [5] H. Avron, S. Toledo, Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix, *J. ACM* 58 (2) (2011) 8.
- [6] E. Bader, M. Kärcher, M.A. Grepl, K. Veroy, Certified reduced basis methods for parametrized distributed elliptic optimal control problems with control constraints, *SIAM J. Sci. Comput.* 38 (6) (2016) A3921–A3946.
- [7] Satish Balay, Shrirang Abhyankar, Mark F. Adams, Jed Brown, Peter Brune, Kris Buschelman, Lisandro Dalcin, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Karl Rupp, Barry F. Smith, Stefano Zampini, Hong Zhang, PETSc users manual. Technical Report ANL-95/11 – Revision 3.6, Argonne National Laboratory, 2015.

- [8] O. Bashir, K. Willcox, O. Ghattas, B. van Bloemen Waanders, J. Hill, Hessian-based model reduction for large-scale systems with initial condition inputs, *Int. J. Numer. Methods Eng.* 73 (2008) 844–868.
- [9] Y. Bazilevs, T.J.R. Hughes, Weak imposition of Dirichlet boundary conditions in fluid mechanics, *Comput. Fluids* 36 (1) (2007) 12–26.
- [10] P. Benner, S. Gugercin, K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems, *SIAM Rev.* 57 (4) (2015) 483–531.
- [11] P. Benner, A. Onwunta, M. Stoll, Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs, *SIAM J. Matrix Anal. Appl.* 37 (2) (2016) 491–518.
- [12] P. Benner, E. Sachs, S. Volkwein, Model order reduction for PDE constrained optimization, in: *Trends in PDE Constrained Optimization*, Springer, 2014, pp. 303–326.
- [13] D. Bertsimas, D.B. Brown, C. Caramanis, Theory and applications of robust optimization, *SIAM Rev.* 53 (3) (2011) 464–501.
- [14] L.T. Biegler, O. Ghattas, M. Heinkenschloss, B. van Bloemen Waanders (Eds.), *Large-Scale PDE-Constrained Optimization*, Lecture Notes in Computational Science and Engineering, vol. 30, Springer-Verlag, Heidelberg, 2003.
- [15] A. Borzi, V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, vol. 8, SIAM, 2011.
- [16] A. Borzi, V. Schulz, C. Schillings, G. Von Winkel, On the treatment of distributed uncertainties in PDE-constrained optimization, *GAMM-Mitt.* 33 (2) (2010) 230–246.
- [17] T. Bui-Thanh, K. Willcox, O. Ghattas, B. van Bloemen Waanders, Goal-oriented, model-constrained optimization for reduction of large-scale systems, *J. Comput. Phys.* 224 (2) (2007) 880–896.
- [18] Tan Bui-Thanh, Carsten Burstedde, Omar Ghattas, James Martin, Georg Stadler, Lucas C. Wilcox, Extreme-scale UQ for Bayesian inverse problems governed by PDEs, in: *SC12: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2012, Gordon Bell Prize finalist.
- [19] Tan Bui-Thanh, Omar Ghattas, Analysis of the Hessian for inverse scattering problems. Part I: inverse shape scattering of acoustic waves, *Inverse Probl.* 28 (5) (2012) 055001.
- [20] Tan Bui-Thanh, Omar Ghattas, Analysis of the Hessian for inverse scattering problems. Part II: inverse medium scattering of acoustic waves, *Inverse Probl.* 28 (5) (2012) 055002.
- [21] Tan Bui-Thanh, Omar Ghattas, Analysis of the Hessian for inverse scattering problems. Part III: inverse medium scattering of electromagnetic waves, *Inverse Probl. Imaging* 7 (4) (2013) 1139–1155.
- [22] Tan Bui-Thanh, Omar Ghattas, A scalable MAP solver for Bayesian inverse problems with Besov priors, *Inverse Probl. Imaging* 9 (1) (2015) 27–54.
- [23] Tan Bui-Thanh, Omar Ghattas, James Martin, Georg 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 (6) (2013) A2494–A2523.
- [24] P. Chen, Sparse quadrature for high-dimensional integration with Gaussian measure, *ESAIM: Math. Model. Numer. Anal.* 52 (2) (2018) 631–657.
- [25] P. Chen, O. Ghattas, Hessian-based sampling for high-dimensional model reduction, *Int. J. Uncertain. Quantificat.* (2019), in press.
- [26] P. Chen, O. Ghattas, Sparse polynomial approximations for optimal control problems constrained by elliptic PDE with lognormal coefficient, preprint, 2019.
- [27] P. Chen, M. Haberman, O. Ghattas, Optimal design of acoustic cloak under uncertainty, preprint, 2019.
- [28] P. Chen, A. Quarteroni, Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraints, *SIAM/ASA J. Uncertain. Quantificat.* 2 (1) (2014) 364–396.
- [29] P. Chen, A. Quarteroni, G. Rozza, Stochastic optimal Robin boundary control problems of advection-dominated elliptic equations, *SIAM J. Numer. Anal.* 51 (5) (2013) 2700–2722.
- [30] P. Chen, A. Quarteroni, G. Rozza, Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations, *Numer. Math.* 133 (1) (2016) 67–102.
- [31] P. Chen, A. Quarteroni, G. Rozza, Reduced basis methods for uncertainty quantification, *SIAM/ASA J. Uncertain. Quantificat.* 5 (1) (2017) 813–869.
- [32] P. Chen, U. Villa, O. Ghattas, Hessian-based adaptive sparse quadrature for infinite-dimensional Bayesian inverse problems, *Comput. Methods Appl. Mech. Eng.* 327 (2017) 147–172.
- [33] Benjamin Crestel, Alen Alexanderian, Georg Stadler, Omar Ghattas, A-optimal encoding weights for nonlinear inverse problems, with application to the Helmholtz inverse problem, *Inverse Probl.* 33 (7) (2017) 074008.
- [34] J. Dick, F.Y. Kuo, I.H. Sloan, High-dimensional integration—the quasi-Monte Carlo way, *Acta Numer.* 22 (2013) 133–288.
- [35] J. Dick, Q.T. Le Gia, Ch. Schwab, Higher order quasi-Monte Carlo integration for holomorphic, parametric operator equations, *SIAM/ASA J. Uncertain. Quantificat.* 4 (1) (2016) 48–79.
- [36] Pearl H. Flath, Lucas C. Wilcox, Volkan Akçelik, Judy Hill, Bart van Bloemen Waanders, Omar Ghattas, Fast algorithms for Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations, *SIAM J. Sci. Comput.* 33 (1) (2011) 407–432.
- [37] D. Ghate, M.B. Giles, Efficient Hessian calculation using automatic differentiation, *AIAA Pap.* 4059 (2007) 2007.
- [38] O. Ghattas, J.-H. Bark, Optimal control of two- and three-dimensional Navier-Stokes flow, *J. Comput. Phys.* 136 (1997) 231–244.
- [39] R. Glowinski, J.L. Lions, *Exact and Approximate Controllability for Distributed Parameter Systems*, Cambridge University Press, Cambridge, UK, 1996.
- [40] M.D. Gunzburger, *Perspectives in Flow Control and Optimization*, vol. 5, SIAM, 2003.
- [41] M.D. Gunzburger, H.C. Lee, J. Lee, Error estimates of stochastic optimal Neumann boundary control problems, *SIAM J. Numer. Anal.* 49 (2011) 1532–1552.
- [42] N. Halko, P.-G. Martinsson, J.A. Tropp, Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions, *SIAM Rev.* 53 (2) (2011) 217–288.
- [43] J.S. Hesthaven, G. Rozza, B. Stamm, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics, 2015.
- [44] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, *Optimization with PDE Constraints*, vol. 23, Springer Science & Business Media, 2008.
- [45] L.S. Hou, J. Lee, H. Manouzi, Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs, *J. Math. Anal. Appl.* 384 (1) (2011) 87–103.
- [46] Tobin Isaac, Noemi Petra, Georg Stadler, Omar Ghattas, Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet, *J. Comput. Phys.* 296 (September 2015) 348–368.
- [47] M. Kärcher, M.A. Grepl, A posteriori error estimation for reduced order solutions of parametrized parabolic optimal control problems, *ESAIM: Math. Model. Numer. Anal.* 48 (6) (2014) 1615–1638.
- [48] M. Klein, A. Sadiki, J. Janicka, Investigation of the influence of the Reynolds number on a plane jet using direct numerical simulation, *Int. J. Heat Fluid Flow* 24 (6) (2003) 785–794.
- [49] D.P. Kouri, D. Heinkenschloss, M. Ridzal, B.G. Van Bloemen Waanders, A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty, *SIAM J. Sci. Comput.* 35 (4) (2012) 1847–1879.
- [50] D.P. Kouri, T.M. Surowiec, Risk-averse PDE-constrained optimization using the conditional value-at-risk, *SIAM J. Optim.* 26 (1) (2016) 365–396.
- [51] Drew P. Kouri, Matthias Heinkenschloss, Denis Ridzal, Bart van Bloemen Waanders, Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty, *SIAM J. Sci. Comput.* 36 (6) (2014) A3011–A3029.

- [52] A. Kunoth, Ch. Schwab, Analytic regularity and GPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs, *SIAM J. Control Optim.* 51 (3) (2013) 2442–2471.
- [53] A. Kunoth, Ch. Schwab, Sparse adaptive tensor Galerkin approximations of stochastic PDE-constrained control problems, *SIAM/ASA J. Uncertain. Quantificat.* 4 (1) (2016) 1034–1059.
- [54] T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza, Boundary control and shape optimization for the robust design of bypass anastomoses under uncertainty, *ESAIM: Math. Model. Numer. Anal.* 47 (4) (2013) 1107–1131.
- [55] Finn Lindgren, Håvard Rue, Johan Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach, *J. R. Stat. Soc., Ser. B, Stat. Methodol.* 73 (4) (2011) 423–498.
- [56] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.
- [57] A. Logg, K.A. Mardal, G. Wells, *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*, vol. 84, Springer Science & Business Media, 2012.
- [58] James Martin, Lucas C. Wilcox, Carsten Burstedde, Omar Ghattas, A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion, *SIAM J. Sci. Comput.* 34 (3) (2012) A1460–A1487.
- [59] F. Negri, A. Manzoni, G. Rozza, Reduced basis approximation of parametrized optimal flow control problems for the Stokes equations, *Comput. Math. Appl.* 69 (4) (2015) 319–336.
- [60] F. Negri, G. Rozza, A. Manzoni, A. Quarteroni, Reduced basis method for parametrized elliptic optimal control problems, *SIAM J. Sci. Comput.* 35 (5) (2013) A2316–A2340.
- [61] L.W.T. Ng, K.E. Willcox, Multifidelity approaches for optimization under uncertainty, *Int. J. Numer. Methods Eng.* 100 (10) (2014) 746–772.
- [62] J. Nocedal, S. Wright, *Numerical Optimization*, Springer Science & Business Media, 2006.
- [63] B. Peherstorfer, K. Willcox, M. Gunzburger, Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *SIAM Rev.* 60 (3) (2018) 550–591.
- [64] Noemi Petra, James Martin, Georg Stadler, Omar Ghattas, A computational framework for infinite-dimensional Bayesian inverse problems: part II. Stochastic Newton MCMC with application to ice sheet inverse problems, *SIAM J. Sci. Comput.* 36 (4) (2014) A1525–A1555.
- [65] A. Quarteroni, A. Manzoni, F. Negri, *Reduced Basis Methods for Partial Differential Equations: An Introduction*, vol. 92, Springer, 2015.
- [66] Farbod Roosta-Khorasani, Uri Ascher, Improved bounds on sample size for implicit matrix trace estimators, *Found. Comput. Math.* 15 (5) (2015) 1187–1212.
- [67] E. Rosseel, G.N. Wells, Optimal control with stochastic PDE constraints and uncertain controls, *Comput. Methods Appl. Mech. Eng.* 213–216 (2012) 152–167.
- [68] G. Rozza, D.B.P. Huynh, A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations, *Arch. Comput. Methods Eng.* 15 (3) (2008) 229–275.
- [69] A.K. Saibaba, A. Alexanderian, I.C.F. Ipsen, Randomized matrix-free trace and log-determinant estimators, *Numer. Math.* (2016) 1–43.
- [70] A.K. Saibaba, J. Lee, P.K. Kitanidis, Randomized algorithms for generalized Hermitian eigenvalue problems with application to computing Karhunen–Loève expansion, *Numer. Linear Algebra Appl.* 23 (2) (2016) 314–339.
- [71] C. Schillings, S. Schmidt, V. Schulz, Efficient shape optimization for certain and uncertain aerodynamic design, *Comput. Fluids* 46 (1) (2011) 78–87.
- [72] C. Schillings, Ch. Schwab, Sparse, adaptive Smolyak quadratures for Bayesian inverse problems, *Inverse Probl.* 29 (6) (2013).
- [73] Alexander Shapiro, Darinka Dentcheva, Andrzej Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, Society for Industrial and Applied Mathematics, 2009.
- [74] H. Tiesler, R.M. Kirby, D. Xiu, T. Preusser, Stochastic collocation for optimal control problems with stochastic PDE constraints, *SIAM J. Control Optim.* 50 (5) (2012) 2659–2682.
- [75] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, vol. 112, American Mathematical Society, Providence, RI, 2010.
- [76] U. Villa, N. Petra, O. Ghattas, hippylib: an extensible software framework for large-scale deterministic and linearized Bayesian inversion, <http://hippylib.github.io>, 2016.
- [77] M.J. Zahr, C. Farhat, Progressive construction of a parametric reduced-order model for PDE-constrained optimization, *Int. J. Numer. Methods Eng.* 102 (5) (2015) 1111–1135.