Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2018-0295

Vol. **27**, No. 3, pp. 753-774
March 2020

A Distributed Optimal Control Problem with Averaged Stochastic Gradient Descent

Qi Sun[1,2,*] and Qiang Du[3]

[1] *School of Mathematical Science, University of Science and Technology of China, Hefei 230026, P.R. China.*
[2] *Division of Applied and Computational Mathematics, Beijing Computational Science Research Center, Beijing 100094, P.R. China.*
[3] *Department of Applied Physics and Applied Mathematics, Columbia University, New York 10027, USA.*

Received 6 November 2018; Accepted (in revised version) 20 March 2019

Abstract. In this work, we study a distributed optimal control problem, in which the governing system is given by second-order elliptic equations with log-normal coefficients. To lessen the curse of dimensionality that originates from the representation of stochastic coefficients, the Monte Carlo finite element method is adopted for numerical discretization where a large number of sampled constraints are involved. For the solution of such a large-scale optimization problem, stochastic gradient descent method is widely used but has slow convergence asymptotically due to its inherent variance. To remedy this problem, we adopt an averaged stochastic gradient descent method which performs stably even with the use of relatively large step sizes and small batch sizes. Numerical experiments are carried out to validate our theoretical findings.

AMS subject classifications: 35R60, 78M31, 49J20

Key words: PDE-constrained elliptic control, high-dimensional random inputs, Monte Carlo finite element, stochastic gradient descent.

1 Introduction

In many engineering applications involving uncertainty in the input data, a robust force control is often designed such that the response of governing system is optimal in some sense [23,27]. These problems can often be formulated as the minimization of an objective functional subject to partial differential equations (PDEs) with high-dimensional random inputs. In this case, to overcome the curse of dimensionality and obtain a reliable numerical prediction, the Monte Carlo (MC) methods are typically used in conjunction with

*Corresponding author. *Email addresses:* `sunqi@csrc.ac.cn` (Q. Sun), `qd2125@columbia.edu` (Q. Du)

the associated Galerkin finite element (FE) approximation in space [12, 24, 33, 43]. However, the simulation of MC FE solutions typically requires a large number of constrained equations that may incur an enormous computational cost. As such, efficient algorithms for solving these large-scale PDE-constrained optimization problems are highly desirable in practice. On the other hand, for problems where the dimension of stochastic space is moderate and the solution has a very smooth dependence on the input random variables, better convergence can be achieved using more sophisticated techniques such as sparse grid stochastic collocation methods [7, 30, 44], reduced basis methods [6, 8, 49], low-rank tensor methods [3, 4, 20], etc.

To be specific, we study in this paper the distributed elliptic optimal control problem with log-normal stochastic coefficients. That is, given a deterministic target function $U: D \to \mathbb{R}$, the control objective is formulated as the minimization of a cost functional

$$\mathcal{J}_\beta(u,f) = \mathbb{E}\left[\frac{1}{2}\int_D |u-U|^2 dx\right] + \frac{\beta}{2}\int_D |f|^2 dx \tag{1.1}$$

subject to an elliptic boundary value problem with stochastic coefficient $a(x,\omega) = e^{g(x,\omega)}$, namely, for almost every (a.e.) ω in a set of outcomes Ω,

$$\begin{cases} -\nabla\cdot(a(x,\omega)\nabla u(x,\omega)) = f(x) & \text{in } D, \\ u(x,\omega) = 0 & \text{on } \partial D. \end{cases} \tag{1.2}$$

Here, and in what follows, D denotes the spatial domain, ∂D its boundary, β a positive regularization parameter, $g(\cdot,\omega) = \log(a(\cdot,\omega))$ a normal distributed random field, f the robust control that belongs to a closed, convex, and nonempty admissible set $\mathcal{A} \subset L^2(D)$, and "$\nabla$" means differentiation with respect to (w.r.t.) the spatial variable $x \in D$.

The stochastic coefficients in (1.2) can be expressed via the Karhumen–Loève (KL) expansion. To define such an expansion, we assume the known information on $g(x,\omega)$ includes its mean value and covariance function. Since the decay rate for KL eigenvalues is directly related to the regularity of covariance kernel function [19], the truncated KL expansion may lead to a high-dimensional problem that suffers from the curse of dimensionality. In contrast to the aforementioned collocation-based approaches, we adopt the MC FE method for the numerical approximation of robust control as depicted in Fig. 1. Moreover, the error estimate associated with approximating f^* by $f^*_{h,n}$ is deduced, which implies that a small mesh size and a large number of sampled constraints are required to obtain a satisfactory accuracy.

To solve the large-scale optimization problem resulted from the MC FE discretization, the standard gradient descent (GD) method requires the update of gradient over all samples, thus demanding repeated and costly solutions of the state and adjoint equations. While the deployment of high performance computing platforms using scalable algorithms can significantly speed up the computation, they may not be feasible for many control problems that demand online feedback and miniature design of control devices.