

EFFICIENT NUMERICAL METHODS FOR ELLIPTIC OPTIMAL CONTROL PROBLEMS WITH RANDOM COEFFICIENT

XIAOWEI PANG¹ AND HAIMING SONG^{1,*}

¹ School of Mathematics, Jilin University
Changchun 130012, China

XIAOSHEN WANG² AND JIACHUAN ZHANG^{3,4}

² Department of Mathematics and Statistics, University of Arkansas at Little Rock
Little Rock 72204, United States

³ Department of Mathematics, Southern University of Science and Technology
Shenzhen 518055, China

⁴ School of Mathematics and Statistics, Wuhan University
Wuhan 430072, China

ABSTRACT. Efficient numerical methods for solving Poisson equation constraint optimal control problems with random coefficient are discussed in this paper. By applying the finite element method and the Monte Carlo approximation, the original optimal control problem is discretized and transformed into an optimization problem. Taking advantage of the separable structures, Algorithm 1 is proposed for solving the problem, where an alternating direction method of multiplier is used. Both computational and storage costs of this algorithm are very high. In order to reduce the computational cost, Algorithm 2 is proposed, where the multi-modes expansion is introduced and applied. Further, for reducing the storage cost, we propose Algorithm 3 based on Algorithm 2. The main idea is that the random term is shifted to the objective functional, which could be computed in advance. Therefore, we only need to solve a deterministic optimization problem, which could reduce all the costs significantly. Moreover, the convergence analyses of the proposed algorithms are established, and numerical simulations are carried out to test the performances of them.

1. Introduction. It is well known that many physical and engineering problems could be described by the optimal control problems (OCPs) with partial differential equation (PDE) or random/stochastic partial differential equation (RPDE/SPDE) constraints, such as the optimal heat source problem, the optimal design of the aircraft, and the weather forecasting problem etc. For the deterministic PDE optimal control problem, there exists plenty research results, and we refer to [15, 21, 26] for more details. However, only few works have been developed for the RPDE/SPDE optimal control problem.

Generally speaking, the numerical methods for RPDE/SPDE constraint optimal control problems are divided into two categories: sample methods and projection methods. For the former, Cao et al. proposed an improved Monte Carlo method

2020 *Mathematics Subject Classification.* Primary: 90C30; Secondary: 65K10, 65C05, 65N30.

Key words and phrases. Optimal control problem, finite element method, Monte Carlo approximation, alternating direction method of multiplier, multi-modes expansion.

The work is supported by NSF grant No.11701210, 11871245, and 11971221.

* Corresponding author: Haiming Song(songhaiming@jlu.edu.cn).

using compression variance technique to deal with optimal control problems under random Burger constraints in 2003 [7]. Based on the gradient error, Kouri et al. combined the trust region method and the stochastic collocation method to propose an adaptive collocation method, which is superior to the traditional Newtonian conjugate gradient method [17]. For the latter, Xiu developed a polynomial chaos expansion for the target functional, which transforms the original SPDE constraint optimal control problem to a nonlinear problem with finite coefficients [25]. However, no theoretical guarantee was given there. In 2011, Gunzburger and his collaborators transformed the SPDE constraint optimal control problem into a KKT nonlinear system, and then proved the existence of the solution of the KKT system. By using Karhunen Loeve (KL) expansion or appropriate orthogonal decomposition, they obtained a series of deterministic optimization systems numerically [13]. In 2013, Kunoth and Schwab used generalized polynomial chaos expansion (gPC) to handle the RPDE/SPDE optimal control problems [18]. Recently, Taylor approximation and variance reduction method was proposed by Chen et al. for PDE-constrained optimal control under uncertainty [8].

In this paper, we mainly concentrate on the efficient numerical methods for the optimal control problem constrained by the Poisson equation with random coefficient. The major challenges for dealing with the randomness are: (I) How to deal with the uncertainties appeared in the constraint and the objective functional, and formulate a discretized optimization problem. (II) How to solve the resulted optimization problem, and propose an efficient algorithm. We will analyze the challenges above one by one, and present the corresponding solutions.

Monte Carlo (MC) method [5, 7], the stochastic collocation (SC) method [1, 2], the stochastic Galerkin (SG) method [22, 23], the polynomial chaos (PC) expansion [16], and so on are some of the commonly used methods for dealing with the first issue. Based on the generality and parallelizability of MC method, we will apply it to deal with the random sample space. Further, using the finite element method (FEM) to approximate the physical space, we shall obtain a discretized optimization problem. This is the most natural idea, but it will face huge difficulties in computation and storage when the high accuracy solution is required. The KL expansion is a spectral representation of the covariance function, which consists of the eigenvalues and eigenfunctions. This method is widely used in RPDE/SPDE fields when the eigenvalues decay exponentially or non-Gaussian spatially-dependent random with large correlation length, otherwise it will run into the curse of dimensionality [3]. As an improvement, the multi-modes expansion (MME) [11] will be considered to approximate the state variable firstly, and then carry out MC and FEM subsequently. By exploiting the relationships among the terms of MME expansion, we can transform the discretized optimization problem into a quadratic programming problem with deterministic linear constraints. In this paper, we will follow this approach to resolve the first challenge.

The second challenge is closed related to the first one. If the first challenge can be handled very well, the latter one will be much less difficult. For solving the resulted optimization problem, there exist fruitful algorithms, such as the gradient descent method, the conjugate gradient method, the Newton's method, the regularization method, and the alternating direction method of multipliers (ADMM) etc., the interested readers are referred to [13, 21] and references therein for the rich literature. Because of the separability of state and control of the resulted optimization

problem, ADMM is applied in this paper, which has been used in the optimal control problem with deterministic PDE constraints, and displayed satisfactory results [14, 26]. Moreover, to the best of our knowledge, there is no complete convergence analysis for elliptic RPDE-constrained optimization problems. The existing results merely present the difference between the exact solution and the discretized solution [13, 19], while the error for solving the optimization problem is ignored. Here, based on the convergence analysis of ADMM in [6] and the techniques for dealing with the former challenge, the global error estimate in terms of the deviation between objective functionals will be established for our proposed algorithms later.

For the clarity of the structure, the efficient algorithm proposed in this paper is developed by three steps. Firstly, the MC, the FEM, and the ADMM are used sequentially to formulate a primitive algorithm, Algorithm 1, which needs to compute $M + 1$ inverses of $N \times N$ matrices, and store $\mathcal{O}(M + 1)$ matrices (see Section 3), where M and N denote the sample number in probability space and the degree of freedom in physical space, respectively. To reduce the computation cost, we propose algorithm 2, where the MME technique is introduced. Using this algorithm, we only need to compute inverse matrix twice, but still need to store $\mathcal{O}(M + 1)$ matrices (see Section 4). As the last step, by applying the recurrence relation on MME, the random term in the constraint is shifted to the objective functional. Then, based on Algorithm 2, we propose Algorithm 3, in which we only need to compute inverse matrix twice, and store $\mathcal{O}(1)$ matrices (see Section 5). Naturally, Algorithm 3 reduces the computation cost and storage cost significantly.

The outline of this paper is as follows. In Section 2, the model problem, the well-posedness of the solution, and the first order optimal condition of the model problem are introduced. Algorithm 1, 2, 3, and their convergence analyses are presented in Section 3, 4, and 5, respectively. In Section 6, numerical simulations are carried out to test the performance of the proposed algorithms. Finally, some conclusion remarks are given in Section 7.

2. Optimal control problem with random coefficient. The model problem, the well-posedness of the studied model, and its corresponding first order optimal condition will be presented in this section.

2.1. The model in concern. For easy of description, most of the notations and definitions used in this paper are adopted from the reference [9]. We consider the model problem on a polygon $D \subset \mathbb{R}^2$ with the boundary ∂D . $L^p(D)$, $L^p(\partial D)$, $H^p(D)$, and $H^p(\partial D)$ denote the Lebesgue integrable function spaces and Hilbert spaces on D and ∂D , respectively ($1 \leq p \leq +\infty$). In addition, we introduce a standard probability space $(\Theta, \mathcal{F}, \mathcal{P})$, and a Bochner space $L^p(\Theta, X)$ with the norms

$$\|v\|_{L^p(\Theta, X)} = \left(\int_{\Theta} \|v(\xi, \cdot)\|_X^p d\mathcal{P}(\xi) \right)^{\frac{1}{p}} < \infty, \quad 1 \leq p < \infty,$$

$$\|v\|_{L^\infty(\Theta, X)} = \operatorname{ess\,sup}_{\xi \in \Theta} \|v(\xi, \cdot)\|_X < \infty,$$

where X stands for a Hilbert space.

Using the notions above, the Poisson equation constraint optimal control problem with random coefficient and Dirichlet boundary condition could be written as

$$(OCP) \quad \begin{cases} \min_{y \in Y, u \in U} F(y(\xi, x), u(x)) \\ \text{subject to :} \\ \begin{cases} -\nabla \cdot (\alpha(\xi, x) \nabla y(\xi, x)) = u(x), & \text{in } D, \\ y(\xi, x) = g(x), & \text{on } \partial D, \end{cases} \end{cases}$$

with

$$F(y(\xi, x), u(x)) := \mathbb{E} \left[\frac{1}{2} \int_D |y(\xi, x) - y_d|^2 dx \right] + \frac{\gamma}{2} \int_D |u(x) - u_d|^2 dx, \tag{1}$$

where $y_d \in L^2(D)$ and $u_d \in H^2(D)$ stand for the desirable state variable and the control variable respectively, which are given in advance. The random diffusion coefficient $\alpha(\xi, x)$ belongs to $L^2(\Theta, D)$, where ξ represents a random sample in Θ . The expectation of $\rho(\xi)$ is defined by $\mathbb{E}[\rho(\xi)] = \int_{\Theta} \rho(\xi) d\mathcal{P}(\xi)$, and the boundary condition $g \in L^2(\partial D)$ is a known function, which is set to be 0 in this paper for simplicity.

The corresponding variational formulation of the Poisson equation with random coefficient is given by: Finding $y \in L^2(\Theta, H_0^1(D))$, such that

$$\mathbb{E}[a(y, v)] = \mathbb{E}[(u, v)_{L^2(D)}], \quad \forall v \in L^2(\Theta, H_0^1(D)), \tag{2}$$

with $a(y, v) = (\alpha(\xi, x) \nabla y, \nabla v)_{L^2(D)}$. In order to ensure the well-posedness of the variational formulation (2), the following assumptions on α are needed.

Assumption 1. ([3]) *The random diffusion coefficient $\alpha(\xi, x)$ in (OCP) is assumed to be a strongly measurable mapping from Θ to $L^\infty(D)$. And there exist constants $0 < \alpha_1 < \alpha_2 < \infty$ such that random coefficient $\alpha(x, \xi)$ is uniformly elliptic, i.e., for every $\xi \in \Theta$, there holds*

$$0 < \alpha_1 \leq \text{ess inf}_D \alpha(\xi, x) \leq \|\alpha(\xi, \cdot)\|_{L^\infty(D)} \leq \alpha_2 < \infty.$$

Lemma 2.1. ([2, 4, 12]) *If Assumption 1 holds, then for any given right-hand-side function $u \in L^2(D)$, the variational equation (2) has a unique solution $y(\xi, x) \in L^2(\Theta, H_0^1(D))$. Furthermore, we have the following pointwise H^1 -norm estimate*

$$\|y(\xi, \cdot)\|_{H^1(D)} \leq \frac{1}{\alpha_1} \|u\|_{H^{-1}(D)} \leq \frac{1}{\alpha_1} \|u\|_{L^2(D)}, \tag{3}$$

and the global L^2 -error estimate

$$\mathbb{E} \left[\|y\|_{L^2(D)}^2 \right] \leq C(\alpha) \|u\|_{L^2(D)}^2, \tag{4}$$

where $C(\alpha)$ is a positive constant depending on α_1 and α_2 .

2.2. First order optimal condition. For simplicity, the diffusion coefficient $\alpha(\xi, x)$ in (OCP) is set to be

$$\alpha(\xi, x) = 1 + \varepsilon \sigma(\xi, x), \tag{5}$$

other cases could be considered similarly. Here, $\sigma(\xi, x) \in L^2(\Theta, L^\infty(D))$ denotes the random process satisfying

$$\mathcal{P}\{\xi \in \Theta; \|\sigma(\xi, x)\|_{L^\infty(D)} \leq 1\} = 1,$$

and ε stands for the magnitude of the perturbation. It is not difficult to obtain that the random coefficient α defined in (5) satisfies Assumption 1. Therefore, according to Lemma 2.1, there exists a linear operator $T : L^2(D) \rightarrow L^2(\Theta, H^1(D))$, such that

$$y = T(u), \tag{6}$$

where u and y satisfy the variational equation (2). Further, it follows from the inequality (3) that the following estimate holds,

$$\mathbb{E} \left[\|T(u)\|_{H^1(D)}^j \right] = \mathbb{E} \left[\|y(u)\|_{H^1(D)}^j \right] \leq C_1 \|u\|_{L^2(D)}^j, \quad (7)$$

where $j = 2$ or 4 , and C_1 is a constant depending on the bound of α .

Therefore, if the state variable space is $Y = L^2(\Theta, H_0^1(D))$, and the control variable space is $U = L^2(D)$, the optimal control problem (OCP) could be rewritten as

$$(P) \quad \begin{cases} \min_{y \in Y, u \in U} F(y(\xi, x), u(x)) \\ \text{subject to } y(\xi, x) = T(u(x)), \end{cases}$$

which can also be reformulated as

$$\begin{aligned} \min_{u \in U} J(u) &= \min_{u \in U} F(T(u), u) \\ &= \min_{u \in U} \left\{ \mathbb{E} \left[\frac{1}{2} \int_D |T(u) - y_d|^2 dx \right] + \frac{\gamma}{2} \int_D |u - u_d|^2 dx \right\}. \end{aligned} \quad (8)$$

Here, J is a functional from $L^2(D)$ to \mathbb{R} , and the existence and uniqueness results of the solution for (8) could be found in [16, 19].

Finally, let u^* be the optimal solution of (8). By applying the definition of the Fréchet derivative [9], the linearity of the expectation operator \mathbb{E} , and the solution operator T defined in (6), the first order optimal condition of (P) can be written as

$$J'(u^*) = \mathbb{E}[T^*(T(u^*) - y_d)] + \gamma(u^* - u_d) = 0, \quad (9)$$

where T^* is the adjoint operator of T .

Now, the optimal control problem (P), and its first order optimal condition have all been presented. In the remainder of the paper, we will propose three algorithms for solving the optimal control problem (P) incrementally, and compare them with each other.

3. FEM-MC-ADMM for the model problem. The first algorithm studied in this paper is presented in this section. The motivation comes from the finite element method (FEM), the Monte Carlo approximation (MC), and the alternating directions method of multipliers (ADMM). We will carry out these techniques sequentially to solve the optimal control problem (P), and present an global error estimate in the form of the deviation between the objective functionals.

3.1. FEM discretization. Let \mathcal{K}_h be a regular triangulation of the spatial domain D , then the linear finite element function space could be defined by

$$V_h = \{v_h \in H_0^1(D) \mid v_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{K}_h\},$$

where \mathbb{P}_1 is the space of polynomials of degree less than or equal to 1. Further, let $\mathbf{R} = \mathbf{R}(x) = (\phi_1(x), \dots, \phi_N(x))$ denotes a vector valued function consisting of all the basis functions of V_h , the state variable $y(\xi, x) \in L^2(\Theta, H^1(D))$ can be approximated by

$$y_h(\xi, x) = \sum_{i=1}^N y_i(\xi) \phi_i(x) = \mathbf{R}\mathbf{y}(\xi), \quad (10)$$

which belongs to $Y_h = L^2(\Theta, V_h)$. Here, the vector $\mathbf{y}(\xi) = (y_1(\xi), \dots, y_N(\xi))^T \in \mathbb{R}^N$ for any fixed $\xi \in \Theta$.

Based on above notations, for a fixed sample $\xi \in \Theta$ and a given control variable $u \in L^2(D)$, the FEM approximation for the constraint equation of (OCP) shall be described as : Finding $y_h(\xi, \cdot) \in V_h$ such that

$$a(y_h(\xi, \cdot), v_h) = (u, v_h)_{L^2(D)}, \quad \forall v_h \in V_h. \quad (11)$$

This means there exists an operator $T_h : L^2(D) \rightarrow L^2(\Theta, V_h) \subset L^2(\Theta, H^1(D))$ such that $y_h = T_h(u)$. It is easy to see that T_h is a linear operator.

Lemma 3.1. *The pathwise FEM discretization (11) satisfies the following stability estimate: for any fixed sample $\xi \in \Theta$,*

$$\|T_h(u)\|_{H^1(D)} = \|y_h(\xi, \cdot)\|_{H^1(D)} \leq c(\alpha)\|u\|_{L^2(D)}, \quad (12)$$

where $c(\alpha)$ depends on α .

The proof, which will be omitted here, follows from the coerciveness and continuity of the bilinear form $a(u, v)$. The interested readers are referred to [15].

Now, we generalize the stability estimate of the deterministic problem to the random coefficient case. It is well known that the FEM approximation of (2) is: Finding $y_h \in L^2(\Theta, V_h)$ such that

$$\mathbb{E}[a(y_h, v_h)] = \mathbb{E}[(u, v_h)_{L^2(D)}], \quad \forall v_h \in L^2(\Theta, V_h). \quad (13)$$

By applying the Assumption 1 and the estimate (12), we can derive:

Lemma 3.2. *For any control variable $u \in L^2(D)$, if y_h is the FEM solution of (13), then we have*

$$\mathbb{E} \left[\|T_h u\|_{H^1(D)}^j \right] = \mathbb{E} \left[\|y_h\|_{H^1(D)}^j \right] \leq C_2 \|u\|_{L^2(D)}^j, \quad (14)$$

where $j = 2, 4$, and C_2 is a positive constant depending on α and j .

By using the estimates (4), (7), (14), the Theorem 4.3 in [10], and the triangle inequality, we can obtain the following lemma.

Lemma 3.3. ([11, 19]) *If the control variable $u \in L^2(D)$, then it holds*

$$\left(\mathbb{E}[\|(T - T_h)u\|_{L^2(D)}^j] \right)^{\frac{1}{j}} \leq C_3 h^2 \|u\|_{L^2(D)}, \quad (15)$$

where $j = 2, 4$, and C_3 is a positive constant independent of the mesh size h .

Similar to (10), the control variable u could be approximated by

$$u_h = \mathbf{R}\mathbf{u} \in U_h = V_h, \quad (16)$$

where $\mathbf{u} = (u_1, \dots, u_N)^T \in \mathbb{R}^N$. Substituting the state variable y and the control variable u in (P) by y_h and u_h , respectively, we can get a stochastic optimization problem

$$(P_{1h}) \quad \begin{cases} \min_{y_h \in Y_h, u_h \in U_h} F(y_h(\xi, x), u_h(x)) \\ \text{subject to } y_h(\xi, x) = T_h(u_h(x)). \end{cases}$$

3.2. MC approximation. In this subsection, we mainly apply the standard MC method to approximate the expectation \mathbb{E} in (P_{1h}) . Let $\{\xi_j\}_{j=1}^M$ ($M \gg 1$) be the identically distributed independent samples chosen from the probability space (Θ, \mathcal{F}, P) . For any state variable $y \in L^2(\Theta, H^1(D))$, by using the law of large numbers and the central limit theorem, we can approximate the expectation $\mathbb{E}[y_h(\xi, x)]$ by

$$\mathbb{E}_M[y_h(\xi, x)] = \frac{1}{M} \sum_{j=1}^M y_h(\xi_j, x), \quad (17)$$

which satisfies the error estimate (cf. [4])

$$\left(\mathbb{E} \left[\|\mathbb{E}[y_h] - \mathbb{E}_M[y_h]\|_{L^2(D)}^2 \right] \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{M}} \left(\mathbb{E}[\|y_h\|_{L^2(D)}^2] \right)^{\frac{1}{2}}. \quad (18)$$

Substituting the MC approximation (17) into (P_{1h}) , we can derive the following discretized optimization problem:

$$(P_{1h}^M) \quad \begin{cases} \min_{y_h \in Y_h, u_h \in U_h} F_{M,h}(y_h(\xi, x), u_h(x)) \\ \text{subject to } y_h(\xi_j, x) = T_h(\xi_j, u_h(x)), \quad j = 1, \dots, M, \end{cases}$$

where

$$F_{M,h}(y_h, u_h) = \mathbb{E}_M \left[\frac{1}{2} \int_D |y_h - y_d|^2 dx \right] + \frac{\gamma}{2} \int_D |u_h - u_d|^2 dx. \quad (19)$$

Similar to (P), the optimization problem (P_{1h}^M) has a unique solution $u_{M,h}^* \in V_h$, which satisfies the first order optimal condition

$$\mathbb{E} \left[\mathbb{E}_M [T_h^*(T_h(u_{M,h}^*) - y_d)] \right] + \gamma(u_{M,h}^* - \mathbb{Q}_h u_d) = 0, \quad (20)$$

where T_h^* is the adjoint operator of T_h , and \mathbb{Q}_h is the L^2 projection operator from $L^2(D)$ to V_h .

Using the first order optimal conditions (9) and (20), we can get

Lemma 3.4. ([19]) *Let u^* and $u_{M,h}^*$ be the optimal controls of (P) and (P_{1h}^M) , respectively, we have the following estimate*

$$\|u^* - u_{M,h}^*\|_{L^2(D)} \leq \frac{K_1}{\sqrt{M}} + K_2 h^2 \quad (21)$$

with

$$K_1 = \frac{C}{\gamma} \sqrt[4]{8} (\|y_d\|_{L^2(D)} + \|u^*\|_{L^2(D)}), \quad K_2 = \frac{C}{\gamma} (\|y_d\|_{L^2(D)} + \|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)}),$$

where the constant C is independent of M and h .

Further, applying the definitions of the functionals $F(y, u)$ and $F_{M,h}(y_h, u_h)$, the operators T and T_h , and the error estimate (21) in Lemma 3.4, we can get

Theorem 3.5. *Let (y^*, u^*) and $(y_{M,h}^*, u_{M,h}^*)$ be the optimal solutions of the optimal control problem (P) and its approximation (P_{1h}^M) , respectively, then we have*

$$\mathbb{E} [|F(y^*, u^*) - F_{M,h}(y_{M,h}^*, u_{M,h}^*)|] \leq \frac{\bar{K}_1}{\sqrt{M}} + \bar{K}_2 h^2, \quad (22)$$

where

$$\begin{aligned}\bar{K}_1 &= \frac{\sqrt{8}}{2} (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})^2 + \gamma(\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)})K_1 \\ &\quad + (K_1 C_1 C_2 \|u^*\|_{L^2(D)} + K_1 C_2 \|y_d\|_{L^2(D)}) + 2\sqrt{2} (C_2 K_1 (C_1 \|u^*\|_{L^2(D)} \\ &\quad + \|y_d\|_{L^2(D)})) + 2\sqrt{2} ((C_1 C_3 \|u^*\|_{L^2(D)} + C_3 \|y_d\|_{L^2(D)} + C_1 C_2 K_2) \|u^*\|_{L^2(D)} \\ &\quad + C_2 K_2 \|y_d\|_{L^2(D)}) + 4\sqrt{2} (C_3^2 \|u^*\|^4 + C_2^2 K_2^2 + 2C_2^2 K_1^2), \\ \bar{K}_2 &= ((C_1 C_3 \|u^*\|_{L^2(D)} + C_3 \|y_d\|_{L^2(D)} + C_1 C_2 K_2) \|u^*\|_{L^2(D)} + C_2 \|y_d\|_{L^2(D)} K_2) \\ &\quad + \gamma(\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)})K_2 + 4\sqrt{2} (C_3^2 \|u^*\|^4_{L^2(D)} + C_2^2 K_2^2).\end{aligned}$$

For the sake of structural consideration, we put the proof in Appendix.

3.3. ADMM algorithm for (\mathbf{P}_{1h}^M) . In this subsection, we will introduce the ADMM to solve the optimization problem (\mathbf{P}_{1h}^M) , and present the error estimate of this algorithm. For these purposes, we need to rewrite the optimization problem (\mathbf{P}_{1h}^M) in the matrix-vector form firstly. Replacing y_h and u_h in (19) by $\mathbf{R}\mathbf{y}(\xi)$ and $\mathbf{R}\mathbf{u}$, which have been defined in (10) and (16), the optimization problem (\mathbf{P}_{1h}^M) can be rewritten as

$$(\tilde{\mathbf{P}}_{1h}^M) \quad \begin{cases} \min_{\mathbf{y}(\xi_j), \mathbf{u} \in \mathbb{R}^N} \mathcal{F}_{M,h}(\mathbf{y}(\xi), \mathbf{u}) \\ \text{subject to } A_1(\xi_j)\mathbf{y}(\xi_j) = B_1\mathbf{u}, \quad j = 1, \dots, M, \end{cases}$$

where $\mathcal{F}_{M,h}(\mathbf{y}(\xi), \mathbf{u}) = F_{M,h}(\mathbf{R}\mathbf{y}(\xi), \mathbf{R}\mathbf{u}) = F_{M,h}(y_h, u_h)$, and

$$A_1(\xi_j) = (\alpha(\xi_j, x)\nabla\phi_m, \nabla\phi_n)_{L^2(D)} \in \mathbb{R}^{N \times N}, \quad B_1 = (\phi_m, \phi_n)_{L^2(D)} \in \mathbb{R}^{N \times N}.$$

The augmented Lagrangian functional of the optimization problem $(\tilde{\mathbf{P}}_{1h}^M)$ is

$$\begin{aligned}\mathcal{L}_{1\beta}(\mathbf{y}(\xi), \mathbf{u}, \lambda(\xi)) \\ = \mathcal{F}_{M,h}(\mathbf{y}(\xi), \mathbf{u}) - \mathbb{E}_M [(\lambda(\xi), A_1(\xi)\mathbf{y}(\xi) - B_1\mathbf{u})] + \mathbb{E}_M \left[\frac{\beta}{2} \|A_1(\xi)\mathbf{y}(\xi) - B_1\mathbf{u}\|^2 \right],\end{aligned}$$

where $\lambda(\xi) \in \mathbb{R}^N$ is the Lagrange multipliers and β denotes the penalty parameter. Here and hereafter, the notation (\cdot, \cdot) represents the inner product of vectors, and $\|\cdot\|$ stands for the l^2 norm of a vector. It is clear that the optimization problem $(\tilde{\mathbf{P}}_{1h}^M)$ equals the saddle-point problem (cf. [21])

$$\min_{\mathbf{y}(\xi), \mathbf{u}} \max_{\lambda(\xi)} \mathcal{L}_{1\beta}(\mathbf{y}(\xi), \mathbf{u}, \lambda(\xi)). \quad (23)$$

For the saddle-point problem (23), the ADMM is an efficient algorithm. We will apply ADMM to solve this problem, and the whole process can be described by

Notice that every iteration of the ADMM has the closed-form solution

$$\begin{aligned}\mathbf{y}^{t+1}(\xi_j) &= (B_1 + \beta A_1(\xi_j)^T A_1(\xi_j))^{-1} \\ &\quad (c + \beta A_1(\xi_j)^T B_1 \mathbf{u}^t + A_1(\xi_j)^T \lambda^t(\xi_j)), \\ \mathbf{u}^{t+1} &= (\gamma B_1 + \beta B_1^T B_1)^{-1} \mathbb{E}_M [\gamma d + \beta B_1^T A_1(\xi) \mathbf{y}^{t+1}(\xi) - B_1^T \lambda^t(\xi)], \\ \lambda^{t+1}(\xi_j) &= \lambda^t(\xi_j) - \beta (A_1(\xi_j) \mathbf{y}^{t+1}(\xi_j) - B_1 \mathbf{u}^{t+1}),\end{aligned} \quad (24)$$

with $j = 1, \dots, M$.

Now, we consider the convergence of the ADMM. It is easy to find that the stiffness matrixs $A_1(\xi_j)$, $j = 1, \dots, M$, and the mass matrix B_1 are all symmetric

Algorithm 1 *FEM-MC-ADMM*.**Begin**

Let $t = 0$, and set the initial values ε_1^0 , \mathbf{u}^0 , $\{\lambda^0(\xi_j)\}_{j=1}^M$;
 Compute $c = (y_d, \phi_n(x))_{L^2(D)} \in \mathbb{R}^N$, $d = (u_d, \phi_n(x))_{L^2(D)} \in \mathbb{R}^N$;
 Generate the mass matrix B_1 ;
for $j = 1, 2, \dots, M$ **do**
 Generate the stiff matrix $A_1(\xi_j)$;
end for
while $\varepsilon_1^t > \varepsilon_1^0$ **do**

$$\begin{aligned} \mathbf{y}^{t+1}(\xi_j) &= \arg \min_{\mathbf{y}(\xi_j)} \mathcal{L}_{1\beta}(\mathbf{y}(\xi_j), \mathbf{u}^t, \lambda^t(\xi_j)); \quad \text{for } j = 1, \dots, M; \\ \mathbf{u}^{t+1} &= \arg \min_{\mathbf{u}} \mathcal{L}_{1\beta}(\mathbf{y}^{t+1}(\xi), \mathbf{u}, \lambda^t(\xi)); \\ \lambda^{t+1}(\xi_j) &= \lambda^t(\xi_j) - \beta(A_1(\xi_j)\mathbf{y}^{t+1}(\xi_j) - B_1\mathbf{u}^{t+1}); \quad \text{for } j = 1, \dots, M; \end{aligned}$$

Increment $t = t + 1$;

Update $\varepsilon_1^t = \|\mathbf{u}^{t+1} - \mathbf{u}^t\| + \|\lambda^{t+1} - \lambda^t\|$;

end while

End

positive definite matrices, and $\mathcal{F}_{M,h}(\mathbf{y}(\xi), \mathbf{u})$ is a convex functional. For the deterministic case under conditions above, the convergence result has been established in [6]. Applying the Assumption 1, it is not hard to obtain the following convergence result for the random case (cf. [20, 24]).

Lemma 3.6. *For the problem (23), after t iterations, the iterate solution of the objective function generated by ADMM converges at the rate $\mathcal{O}(\frac{1}{t})$, that is*

$$\mathbb{E} [|\mathcal{F}_{M,h}(\mathbf{y}^t, \mathbf{u}^t) - \mathcal{F}_{M,h}(\mathbf{y}_M^*, \mathbf{u}_M^*)|] \leq \frac{\bar{K}_3}{t}, \quad (25)$$

where $(\mathbf{y}_M^*, \mathbf{u}_M^*)$ is the theoretical solution of (23), and

$$\bar{K}_3 = \frac{1}{2} \mathbb{E} \left[\mathbb{E}_M \left[\left\| \begin{array}{c} \mathbf{u}^0 - \mathbf{u}_M^* \\ \lambda^0 - \lambda^* \end{array} \right\|_{\tilde{H}_\beta} + \frac{1}{\beta} \|\lambda^0\| \right] \right], \tilde{H}_\beta = \begin{pmatrix} \beta B_1^T B_1 & 0 \\ 0 & \frac{1}{2\beta} I \end{pmatrix}.$$

Further, applying the Theorem 3.5, the identity

$$F_{M,h}(y_{M,h}^*, u_{M,h}^*) - F_{M,h}(\mathbf{Ry}^t, \mathbf{Ru}^t) = \mathcal{F}_{M,h}(\mathbf{y}_M^*, \mathbf{u}_M^*) - \mathcal{F}_{M,h}(\mathbf{y}^t, \mathbf{u}^t),$$

the Lemma 3.6, and the triangle inequality, we can get the global error estimate.

Theorem 3.7. *A global error estimate for Algorithm 1 is as follows:*

$$\mathbb{E} [|F(y^*, u^*) - F_{M,h}(\mathbf{Ry}^t, \mathbf{Ru}^t)|] \leq \frac{\bar{K}_1}{\sqrt{M}} + \bar{K}_2 h^2 + \frac{\bar{K}_3}{t}. \quad (26)$$

From Algorithm 1, we can see that the stiffness matrix $A_1(\xi)$ is a random one, which needs to be computed and stored for every sample ξ_j , $j = 1, \dots, M$. The cost to compute $A_1(\xi)$ and the mass matrix B_1 is $\mathcal{O}((M+1)N^2)$. In addition, from the closed form solutions (24), we also need to calculate and save the inverse matrix $(\gamma B_1 + \beta B_1^T B_1)^{-1}$, and the inverse matrices $(B_1 + \beta A_1(\xi_j)^T A_1(\xi_j))^{-1}$ for all the samples. Thus, we must compute $M+1$ inverses of $N \times N$ matrices, which will cost a large amount of memory and computation time.

4. MME-FEM-MC-ADMM for the model problem. To reduce the computational cost in Algorithm 1, we shall introduce the multi-modes expansion (MME), and propose a new algorithm. The main idea is to approximate the state variable by MME technique, and then transform the original constraint in Algorithm 1 and thus shift the randomness to the right hand side mass matrix. The advantage is that we can calculate the expectation of the right hand side random matrix in advance, then compute and store the inverse of the matrix only once, which could reduce the computational cost significantly.

4.1. Multi-modes expansion. The multi-modes expansion for the state variable y can be expressed as (cf. [11])

$$y = \sum_{q=0}^{\infty} \varepsilon^q y_q, \quad (27)$$

where ε is the magnitude of the perturbation defined in (5), and $y_q \in L^2(\Theta, H_0^1(D))$. Substituting (27) into the state equation of the optimal control problem (OCP), and comparing the coefficients of the powers of ε , we could get the following Poisson equations in recursive form.

$$\begin{aligned} -\Delta y_0 &= u, & \text{in } D, \\ -\Delta y_q &= \nabla \cdot (\sigma(\xi, \cdot) \nabla y_{q-1}), & \text{in } D, \quad \forall q \geq 1, \\ y_q &= 0, & \text{on } \partial D, \quad \forall q \geq 0. \end{aligned} \quad (28)$$

It is apparent that the relationship between y_0 and u is deterministic, so we can solve the equations (28) by recursion for $q \geq 1$. The well-posedness of the MME solutions has been established in Theorem 3.1 of the reference [11].

Lemma 4.1. *The equations (28) has a unique solution $y_q \in L^2(\Theta, H_0^1(D))$. Besides, for any $q \geq 0$, if $y_q \in L^2(\Theta, H^2(D))$, then the following energy estimate holds*

$$\mathbb{E} \left[\|y_q\|_{H^2(D)}^2 \right] \leq c(\alpha)^{q+1} \|u\|_{L^2(D)}^2, \quad (29)$$

where $c(\alpha)$ is given by (12) and is independent of q and ε .

For the purpose of numerical solution, we truncate the expression (27) after the first Q terms, and denote it by

$$y^Q = \sum_{q=0}^{Q-1} \varepsilon^q y_q. \quad (30)$$

Lemma 4.2. ([11]) *If the perturbation magnitude $\varepsilon < \min\{1, c(\alpha)^{-\frac{1}{2}}\}$, then we have*

$$\lim_{Q \rightarrow \infty} \mathbb{E} \|y - y^Q\|_{L^2(D)} = 0.$$

Lemma 4.1 shows that the problem (28) is well-posed, and Lemma 4.2 shows that the expression (30) is an approximation of the state variable y . Based on these facts, there must exist a linear operator $T^Q : L^2(D) \rightarrow L^2(\Theta, H^1(D))$ such that

$$y^Q = T^Q(u). \quad (31)$$

Feng et al. [11] showed that the truncation error of MME after Q terms is $\mathcal{O}(\varepsilon^Q)$.

Lemma 4.3. *Under the assumptions of Lemma 4.1 and Lemma 4.2, it holds that*

$$\left(\mathbb{E} \left[\|(T - T^Q)u\|_{L^2(D)}^j \right] \right)^{\frac{1}{j}} = \left(\mathbb{E} \left[\|y - y^Q\|_{L^2(D)}^j \right] \right)^{\frac{1}{j}} \leq C_4 \varepsilon^Q \|u\|_{L^2(D)}, \quad (32)$$

where $j = 2$ or 4 , C_4 is a positive constant depending on α , Q and j .

4.2. FEM and MC approximation. In this subsection, we will apply FEM to discretize $y_q (q \geq 0)$ on physical space, and MC method to approximate the expectation \mathbb{E} on probability space. Then the discretized optimization problem, its first order optimal condition, and the convergence analysis for the proposed approach will be presented sequentially.

For any fixed $\xi \in \Theta$, the FEM approximation of $y_q(\xi, x)$ in (28) can be written as

$$y_{q,h}(\xi, x) = \sum_{i=1}^N y_{q,i}(\xi) \phi_i(x) = \mathbf{R} \mathbf{y}_q(\xi), \quad q \geq 0, \quad (33)$$

where $\mathbf{y}_q(\xi) \in \mathbb{R}^N$. Naturally, the truncated MME state $y^Q(\xi, x)$ could be approximated by

$$y_h^Q(\xi, x) = \sum_{q=0}^{Q-1} \varepsilon^q y_{q,h}(\xi, x) = \mathbf{R} \sum_{q=0}^{Q-1} \varepsilon^q \mathbf{y}_q(\xi) = \mathbf{R} \mathbf{y}^Q(\xi). \quad (34)$$

And the variational formulation of the constraint equation in the truncated MME form is given by: For each $\xi \in \Theta$, find $y_h^Q(\xi, x) \in L^2(\Theta, V_h)$ such that

$$\begin{aligned} (\nabla y_{0,h}(x), \nabla v_h(x))_{L^2(D)} &= (u(x), v_h(x))_{L^2(D)}, \\ (\nabla y_{q,h}(\xi, x), \nabla v_h(x))_{L^2(D)} &= (\sigma(\xi, x) \nabla y_{q-1,h}(\xi, x), \nabla v_h(x))_{L^2(D)}, \end{aligned} \quad (35)$$

for any $v_h(x) \in V_h$, where $q = 1, \dots, Q-1$. Similar to (31), there exists a linear operator $T_h^Q : L^2(D) \rightarrow L^2(\Theta, V_h) \subset L^2(\Theta, H^1(D))$ such that

$$y_h^Q = T_h^Q(u). \quad (36)$$

Lemma 4.4. (cf. [11]) *Under the assumption of Lemma 4.2, for $j = 2, 4$ we have the following stability estimate*

$$\mathbb{E} \left[\|T_h^Q u\|_{L^2(D)}^j \right] = \mathbb{E} \left[\|y_h^Q(u)\|_{L^2(D)}^j \right] \leq C_5 \|u\|_{L^2(D)}^j, \quad (37)$$

and the error estimate

$$\begin{aligned} &\left(\mathbb{E} \left[\|(T^Q - T_h^Q)u\|_{L^2(D)}^j \right] \right)^{\frac{1}{j}} \\ &= \left(\mathbb{E} \left[\|y^Q - y_h^Q\|_{L^2(D)}^j \right] \right)^{\frac{1}{j}} \leq C_6 h^2 \|u\|_{L^2(D)}, \end{aligned} \quad (38)$$

where C_5 and C_6 are constants depending on ε and Q .

Further, applying the triangle inequality, the estimates (32) and (38), we have the following error estimate.

Lemma 4.5. *Under the conditions of Lemma 4.1 and Lemma 4.4, we have the following estimate*

$$\left(\mathbb{E} \left[\|(T - T_h^Q)u\|_{L^2(D)}^j \right] \right)^{\frac{1}{j}} \leq C(\varepsilon^Q + h^2) \|u\|_{L^2(D)}, \quad (39)$$

where $j = 2$ or 4 .

By arguments similar to the ones used to get the optimization problem (19), applying the expressions (16), (34) and (36), the original problem (P) could be approximated by

$$(P_{2h}^Q) \quad \begin{cases} \min_{y_h^Q(\xi) \in Y_h, u_h \in U_h} F_{M,h}(y_h^Q(\xi, x), u_h) \\ \text{subject to } y_h^Q(\xi_j, x) = (T_h^Q(u_h))(\xi, x_j) \quad j = 1, \dots, M, \end{cases}$$

where $F_{M,h}$ has been given in (19). Similar to (P_{1h}^M) , the optimization problem (P_{2h}^Q) has a unique solution $u_{M,h}^{Q*} \in V_h$, which satisfies the first order optimal condition

$$\mathbb{E} \left[\mathbb{E}_M \left[T_h^{Q*} (T_h^Q u_{M,h}^{Q*} - y_d) \right] \right] + \gamma (u_{M,h}^{Q*} - \mathbb{Q}_h u_d) = 0. \quad (40)$$

Using the first order optimal conditions (9) and (40), we have

Lemma 4.6. (cf. [19]) *Under the conditions of Lemma 4.1-4.5, the following error estimate*

$$\|u^* - u_{M,h}^{Q*}\|_{L^2(D)} \leq \frac{\tilde{C}_1}{\sqrt{M}} + \tilde{C}_2 \varepsilon^Q + \tilde{C}_3 h^2, \quad (41)$$

holds, where

$$\begin{aligned} \tilde{C}_1 &= \frac{\sqrt[4]{8}}{\gamma} C_5 (C_5 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)}), \\ \tilde{C}_2 &= \frac{CC_4}{\gamma} (\|y_d\|_{L^2(D)} + (C_1 + C_5) \|u^*\|_{L^2(D)}), \\ \tilde{C}_3 &= C \|u_0\|_{L^2(D)} + \frac{1}{\gamma} CC_5 (\|y_d\|_{L^2(D)} + (C_1 + C_5) \|u^*\|_{L^2(D)}). \end{aligned}$$

Similar to Theorem 3.5, we can get the error estimate on the difference of the objective functionals, as follows

Theorem 4.7. *Let (y^*, u^*) and $(y_{M,h}^{Q*}, u_{M,h}^{Q*})$ be the optimal solutions of the optimal control problem (P) and its approximation (P_{2h}^Q) , respectively, then we have*

$$\mathbb{E} \left[|F(y^*, u^*) - F_{M,h}(y_{M,h}^{Q*}, u_{M,h}^{Q*})| \right] \leq \frac{\bar{C}_1}{\sqrt{M}} + \bar{C}_2 \varepsilon^Q + \bar{C}_3 h^2, \quad (42)$$

where

$$\begin{aligned} \bar{C}_1 &= \gamma (\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)}) \tilde{C}_1 + \frac{\sqrt{8}}{2} (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})^2 \\ &\quad + (C_1 C_5 \|u^*\|_{L^2(D)} + C_5 \|y_d\|_{L^2(D)}) \tilde{C}_1 + 2\sqrt{2} (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)}) C_2 \tilde{C}_1, \\ \bar{C}_2 &= \gamma (\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)}) \tilde{C}_2 + ((C_1 C \|u^*\|_{L^2(D)} + C_3 \|y_d\|_{L^2(D)}) \|u^*\|_{L^2(D)} \\ &\quad + 2\sqrt{2} (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)}) C_2 \tilde{C}_2 + (C_1 C_5 \|u^*\|_{L^2(D)} + C_5 \|y_d\|_{L^2(D)}) \tilde{C}_2, \\ \bar{C}_3 &= \gamma (\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)}) \tilde{C}_3 + ((C_1 C \|u^*\|_{L^2(D)} + C_3 \|y_d\|_{L^2(D)}) \|u^*\|_{L^2(D)} \\ &\quad + (C_1 C_5 \|u^*\|_{L^2(D)} + C_5 \|y_d\|_{L^2(D)}) \tilde{C}_3) + 4\sqrt{2} \tilde{C}_3^2. \end{aligned}$$

The proof is similar to Theorem 3.5, we omit it here.

4.3. **ADMM algorithm for (\mathbf{P}_{2h}^Q) .** In this subsection, the ADMM algorithm for solving the optimization problem (\mathbf{P}_{2h}^Q) will be presented. In order to describe the algorithm clearly, we simplify (\mathbf{P}_{2h}^Q) firstly.

Substituting the control variable u in the variational formulation (35) by u_h defined in (16), and using the definition (33) on $y_{q,h}$, the constraint of the optimization problem (\mathbf{P}_{2h}^Q) could be rewritten as

$$A_2 \mathbf{y}_0 = B_1 \mathbf{u}, \quad A_2 \mathbf{y}_q(\xi) = B_0(\xi) \mathbf{y}_{q-1}(\xi), \quad \forall q \geq 1, \quad (43)$$

with

$$A_2 = (\nabla \phi_m, \nabla \phi_n)_{L^2(D)}, \quad B_1 = (\phi_m, \phi_n)_{L^2(D)}, \quad \text{and} \quad B_0(\xi) = (\sigma(\xi) \nabla \phi_m, \nabla \phi_n)_{L^2(D)}.$$

Applying the definition of \mathbf{y}_h^Q in (34), we know that

$$A_2 \mathbf{y}^Q(\xi) = \sum_{q=0}^{Q-1} \varepsilon^q A_2 \mathbf{y}_q(\xi).$$

Further, combining above equality with the equations (43), the relation between state variable $\mathbf{y}^Q(\xi)$ and control variable \mathbf{u} can be given by

$$A_2 \mathbf{y}^Q(\xi) = B_2(\xi) \mathbf{u},$$

where

$$B_2(\xi) = \begin{cases} B_1, & Q = 1, \\ (I + \varepsilon B_0(\xi) A_2^{-1}) B_1, & Q = 2, \\ (I + (\varepsilon B_0(\xi))(I + \varepsilon A_2^{-1} B_0(\xi)) A_2^{-1}) B_1, & Q = 3. \end{cases}$$

Therefore, the optimization problem (\mathbf{P}_{2h}^Q) can be reformulated as:

$$(\tilde{\mathbf{P}}_{2h}^Q) \quad \begin{cases} \min_{\mathbf{y}^Q(\xi), \mathbf{u} \in \mathbb{R}^N} \mathcal{F}_{M,h}(\mathbf{y}^Q(\xi), \mathbf{u}) \\ \text{subject to} \quad A_2 \mathbf{y}^Q(\xi_j) = B_2(\xi_j) \mathbf{u}, \quad j = 1, \dots, M, \end{cases}$$

where the functional $\mathcal{F}_{M,h}$ is defined in (23).

The optimization problem $(\tilde{\mathbf{P}}_{2h}^Q)$ is equivalent to the saddle point problem

$$\min_{\mathbf{y}^Q(\xi), \mathbf{u}} \max_{\lambda(\xi)} \mathcal{L}_{2\beta}(\mathbf{y}^Q(\xi), \mathbf{u}, \lambda(\xi)), \quad (44)$$

with the augmented Lagrangian function

$$\begin{aligned} & \mathcal{L}_{2\beta}(\mathbf{y}^Q(\xi), \mathbf{u}, \lambda(\xi)) \\ &= \mathcal{F}_{M,h}(\mathbf{y}^Q(\xi), \mathbf{u}) - \mathbb{E}_M [(\lambda(\xi), A_2 \mathbf{y}^Q(\xi) - B_2(\xi) \mathbf{u})] \\ &+ \mathbb{E}_M \left[\frac{\beta}{2} \|A_2 \mathbf{y}^Q(\xi) - B_2(\xi) \mathbf{u}\|^2 \right], \end{aligned}$$

where $\lambda(\xi) \in \mathbb{R}^N$ is the Lagrange multipliers and β denotes the penalty parameter.

Now, the ADMM algorithm for solving the saddle point problem (44) could be described by

The ADMM process has the following closed form solutions

$$\begin{aligned} \mathbf{y}^{Q,t+1}(\xi_j) &= (B_1 + \beta A_2^T A_2)^{-1} (c + \beta A_2^T B_2(\xi_j) \mathbf{u}^t + A_2^T \lambda^t(\xi_j)), \\ \mathbf{u}^{t+1} &= (\gamma B_1 + \beta \mathbb{E}_M [B_2(\xi)^T B_2(\xi)])^{-1} \\ &\quad \mathbb{E}_M [\gamma d + B_2(\xi)^T (\beta A_2 \mathbf{y}^{Q,t+1}(\xi) - \lambda^t(\xi))], \\ \lambda^{t+1}(\xi_j) &= \lambda^t(\xi_j) - \beta (A_2 \mathbf{y}^{Q,t+1}(\xi_j) - B_2(\xi) \mathbf{u}^{t+1}), \end{aligned} \quad (45)$$

Algorithm 2 *MME-FEM-MC-ADMM*.**Begin**Let $t = 0$, and set the initial values $\varepsilon_2^0, \mathbf{u}^0, \{\lambda^0(\xi_j)\}_{j=1}^M$;Generate the vectors c, d , and the matrices A_2, B_1 ;**for** $j = 1, 2, \dots, M$ **do** Generate the matrixs $B_0(\xi)$ and $B_2(\xi)$ for the given Q ;**end for****while** $\varepsilon_2^t > \varepsilon_2^0$ **do**

$$\mathbf{y}^{Q,t+1}(\xi_j) = \arg \min_{\mathbf{y}^{Q}(\xi_j)} \mathcal{L}_{2\beta}(\mathbf{y}^Q(\xi_j), \mathbf{u}^t, \lambda^t(\xi_j)), \quad \text{for } j = 1, \dots, M;$$

$$\mathbf{u}^{t+1} = \arg \min_{\mathbf{u}} \mathcal{L}_{2\beta}(\mathbf{y}^{Q,t+1}(\xi), \mathbf{u}, \lambda^t(\xi));$$

$$\lambda^{t+1}(\xi_j) = \lambda^t(\xi_j) - \beta(A_2 \mathbf{y}^{Q,t+1}(\xi_j) - B_2(\xi_j) \mathbf{u}^{t+1}), \quad \text{for } j = 1, \dots, M;$$

 Increment $t = t + 1$; Update $\varepsilon_2^t = \|\mathbf{u}^{t+1} - \mathbf{u}^t\| + \|\lambda^{t+1} - \lambda^t\|$;**end while****End**

with $j = 1, 2, \dots, M$. Similar to Lemma 3.6, let $(\mathbf{y}_M^{Q,*}, \mathbf{u}_M^{Q,*})$ be the exact solution of the problem (\tilde{P}_{2h}^Q) , which satisfies $(y_{M,h}^{Q,*}, u_{M,h}^{Q,*}) = (\mathbf{Ry}_M^{Q,*}, \mathbf{Ru}_M^{Q,*})$, then there exists a constant \tilde{K}_3 such that

$$\begin{aligned} & \mathbb{E} \left[|F_{M,h}(y_{M,h}^{Q,*}, u_{M,h}^{Q,*}) - F_{M,h}(\mathbf{Ry}^{Q,t}, \mathbf{Ru}^t)| \right] \\ &= \mathbb{E} \left[|\mathcal{F}_{M,h}(\mathbf{y}_M^{Q,*}, \mathbf{u}_M^{Q,*}) - \mathcal{F}_{M,h}(\mathbf{y}^{Q,t}, \mathbf{u}^t)| \right] \leq \tilde{K}_3 \frac{1}{t}. \end{aligned}$$

Furthermore, applying the estimate of Theorem 4.7 and triangle inequality, we obtain the following estimate:

Theorem 4.8. *Let $(\mathbf{y}^{Q,t}, \mathbf{u}^t)$ denote the solution obtained after using FEM, MME, and MC approximation and (y^*, u^*) be the solution of (1). Then we have*

$$\mathbb{E} \left[|F(y^*, u^*) - F_{M,h}(\mathbf{Ry}^{Q,t}, \mathbf{Ru}^t)| \right] \leq \frac{\bar{C}_1}{\sqrt{M}} + \bar{C}_2 \varepsilon^Q + \bar{C}_3 h^2 + \tilde{K}_3 \frac{1}{t}, \quad (46)$$

where $F(\cdot, \cdot)$ and $F_{M,h}(\cdot, \cdot)$ are functionals defined by (1) and (19), respectively.

From the iterations (45), we can see that all the matrices are deterministic except the matrix $B_2(\xi)$. Fortunately, the matrix $\mathbb{E}_M[B_2(\xi)^T B_2(\xi)]$ could be obtained in advance for the given samples. Therefore, we need to compute only two inverse matrices, which could reduce the computational cost on the inverse matrices in the Algorithm 1 greatly. However, in addition to the cost of computing and storing A_2 , we still need to compute and store the mass matrices $B_2(\xi_j)$ for all samples. Thus, the storage requirement does not decrease, it is still $\mathcal{O}((M+1)N^2)$.

5. SMME-FEM-MC-ADMM for the model problem. In this section, a new algorithm is proposed to overcome the disadvantages of the Algorithm 2, which can reduce the storage requirement significantly.

From the variational formulation (35), we know that y_q depends on the preceding term y_{q-1} for $q \geq 1$. Therefore, we can establish a relationship between y_0 and y_q .

Furthermore, applying the definition (30) of y^Q , there must exist an relation between y_0 and y^Q . Based on above facts, using the equations (34) and (43), we can obtain

$$\mathbf{y}^Q(\xi) = \sum_{q=0}^{Q-1} \varepsilon^q \mathbf{y}_q(\xi) = P(\xi) \mathbf{y}_0, \quad (47)$$

where $P(\xi)$ is a matrix with random variables, defined as follows

$$P(\xi) = \begin{cases} I, & Q = 1, \\ I + \varepsilon A_2^{-1} B_0(\xi), & Q = 2, \\ I + \varepsilon A_2^{-1} B_0(\xi) + \varepsilon^2 (A_2^{-1} B_0(\xi))^2, & Q = 3. \end{cases}$$

The relationship (47) implies that $\mathbf{y}^Q(\xi)$ is determined by \mathbf{y}_0 , so we need to consider only the constraint on \mathbf{y}_0 . The optimization problem ($\tilde{\mathbf{P}}_{2h}^Q$) shall be simplified as

$$(\tilde{\mathbf{P}}_{3h}^Q) \quad \begin{cases} \min_{\mathbf{y}_0, \mathbf{u} \in \mathbb{R}^N} \mathcal{F}_{M,h}(P(\xi) \mathbf{y}_0, \mathbf{u}) \\ \text{subject to } A_3 \mathbf{y}_0 - B_3 \mathbf{u} = 0, \end{cases}$$

where the functional $\mathcal{F}_{M,h}$ is defined in (23), $A_3 = A_2$ and $B_3 = B_1$ are two deterministic matrices defined in (43). This problem is a classical optimization problem, which is the same as

$$\min_{\mathbf{y}_0, \mathbf{u}} \max_{\lambda} \mathcal{L}_{3\beta}(\mathbf{y}_0, \mathbf{u}, \lambda), \quad (48)$$

with the augmented Lagrangian function

$$\mathcal{L}_{3\beta}(\mathbf{y}_0, \mathbf{u}, \lambda) = \mathcal{F}_{M,h}(P(\xi) \mathbf{y}_0, \mathbf{u}) - (\lambda, A_3 \mathbf{y}_0 - B_3 \mathbf{u}) + \frac{\beta}{2} \|A_3 \mathbf{y}_0 - B_3 \mathbf{u}\|^2,$$

where $\lambda \in \mathbb{R}^N$ is Lagrange multiplier independent of the samples, and β is the penalty parameter.

Now, we present the ADMM algorithm to solve the saddle point problem (48).

Algorithm 3 *SMME-FEM-MC-ADMM*.

Begin

Let $t = 0$, and set the initial values $\varepsilon_3^0, \mathbf{u}^0, \lambda^0$;

Compute the vectors c and d , the matrices A_3 and B_3 ;

for $j = 1, 2, \dots, M$ **do**

Generate the matrices $B_0(\xi_j)$ and $P(\xi_j)$;

end for

while $\varepsilon_3^t > \varepsilon_3^0$ **do**

$$\mathbf{y}_0^{t+1} = \arg \min_{\mathbf{y}_0} \mathcal{L}_{3\beta}(\mathbf{y}_0, \mathbf{u}^t, \lambda^t);$$

$$\mathbf{u}^{t+1} = \arg \min_{\mathbf{u}} \mathcal{L}_{3\beta}(\mathbf{y}_0^{t+1}, \mathbf{u}, \lambda^t);$$

$$\lambda^{t+1} = \lambda^t - \beta(A_3 \mathbf{y}_0^{t+1} - B_3 \mathbf{u}^{t+1});$$

Increment $t = t + 1$;

Update $\varepsilon_3^t = \|\mathbf{u}^{t+1} - \mathbf{u}^t\| + \|\lambda^{t+1} - \lambda^t\|$;

end while

End

By straightforward calculations, we can get

$$\begin{aligned} \mathbf{y}_0^{t+1} &= (\mathbb{E}_M [P(\xi)^T B_3 P(\xi)] + \beta A_3^T A_3)^{-1} \\ &\quad (\mathbb{E}_M [P(\xi)^T c] + \beta A_3^T B_3 \mathbf{u}^t + A_3^T \lambda^t), \\ \mathbf{u}^{t+1} &= (\gamma B_3 + \beta B_3^T B_3)^{-1} (\gamma d + \beta B_3^T A_3 \mathbf{y}_0^{t+1} - B_3^T \lambda^t), \\ \lambda^{t+1} &= \lambda^t - \beta (A_3 \mathbf{y}_0^{t+1} - B_3 \mathbf{u}^{t+1}). \end{aligned} \tag{49}$$

It is not hard to see that Algorithm 3 has the same convergence property as Algorithm 2 (see Theorem 4.8). From the iterations (49), we can see that we don't need to store the information for all the samples. The only matrices needing to be remembered are $\mathbb{E}_M [P(\xi)]$, A_3 , and B_3 , which means the memory cost is $\mathcal{O}(N^2)$. Moreover, we need to compute inverse matrices only twice. The above advantages are in sharp contrast with the Algorithm 1 and Algorithm 2. Hence, Algorithm 3 is the most efficient one both in computational and storage aspects.

6. Numerical experiments. In this section, results of numerical experiments, which are carried out to test the performance of the proposed algorithms, will be reported. The convergence order of the FEM for all algorithms is verified firstly. And then, the ADMM convergence order in the form of the objective function is shown. Finally, to illustrate the advantage of Algorithm 3, we present the storage costs of the three algorithms. All the simulations are conducted using Matlab on a high-performance computer with 16 kernels and 256 GB RAM.

We consider the optimal control problem on the spatial domain $D = [-1, 1]^2$, and assume the random variable $\sigma(x, \cdot)$ associated with the random coefficient $\alpha(\xi, x)$ obeys the uniform distribution on the interval of $[-1, 1]$ for any fixed $x \in D$. The weighted parameter γ of the objective functional is set to be 1, the desired state and control are set to be

$$y_d = \sin(\pi x) \sin(\pi y), \quad u_d = 2\pi^2 \sin(\pi x) \sin(\pi y).$$

It is not difficult to find that y_d and u_d are the optimal state and control of the original problem (P) in the expected sense. In addition, the parameters associated with the ADMM are set to be $\beta = 1$, $\varepsilon_i^0 = 10^{-4}$, $i = 1, 2, 3$, respectively.

To verify the convergence order of the FEM, we fix other parameters firstly. For simplicity, the sample number of the MC, the iteration steps of the ADMM, the magnitude of the perturbation defined in (5), and the truncated length of the MME are set to be $M = 1000$, $t = 20$, $\varepsilon = 10^{-4}$, and $Q = 1$, respectively. Other cases could be discussed in similar manners. Now, we carry out our algorithms on the nested meshes with $h = (1/2)^i$, $i = 1, 2, 3, 4, 5$, and present the results in Figure 1. We emphasize that Algorithm 1 is out of memory even with spatial meshsize $h = 1/32$. For aesthetics, we extend the line to the fifth point. From Figure 1, we can see that the FEM convergence order with respect to the control variable u and the state variable y are all of order 2 for the proposed three algorithms, which coincide with the theoretical analysis in Sections 3-5. To illustrate the correctness of the algorithms intuitively, the images of the optimal control in theory, and the numerical solutions on the proposed algorithms with $h = 1/32$ are shown in Figure 2, where the numerical solution for Algorithm 1 is obtained with $h = 1/16$ for the memory limit.

Now, we will test the performance of ADMM. For this purpose, other parameters in the proposed algorithms need to be fixed in advance. Without loss of generality, we set $M = 1000$, $h = 1/32$, and $Q = 1$ for the latter two algorithms and the

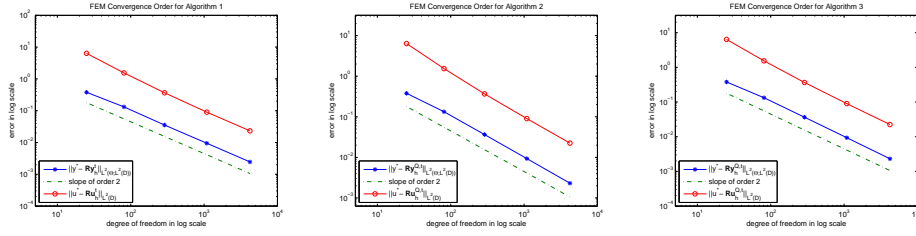


FIGURE 1. The FEM convergence order on the state and the control for three algorithms.

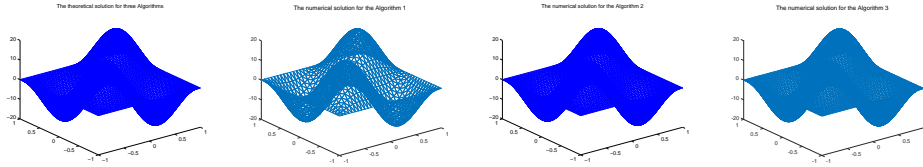


FIGURE 2. The theoretical solution and numerical solutions on the control for three algorithms.

Algorithm 1 with $h = 1/16$. The ADMM errors for the proposed algorithms in the form of the deviation between functionals are shown in Figure 3. From the results, we can see that the convergence rate is superior to $\mathcal{O}(\frac{1}{t})$, which means that the numerical performance of ADMM is better than the theoretical case.

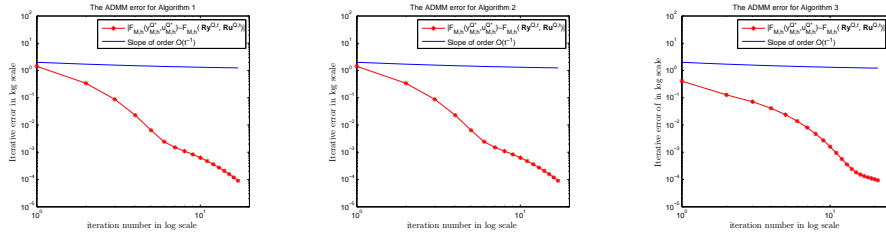


FIGURE 3. The ADMM convergence rate for three algorithms.

Finally, we analyze the CPU memory cost of the proposed algorithms, which is one of the main concerns in this paper. For the fixed MC parameter $M = 1000$, the MME perturbation magnitude $\varepsilon = 10^{-4}$ and the iteration number $t = 20$, the variation trends of the CPU memory cost with respect to the FEM parameter h and the MME parameter Q are given in Table 1. From Table 1, we can find three obvious facts: (1) For any one of the three algorithms, the memory cost increase as the parameter h decreases. (2) For any one of the Algorithm 2 and 3, the memory cost does not have obvious change with respect to the parameter Q . (3) For any fixed h and Q , the memory cost of Algorithm 3 is significantly lower than other two algorithms. The above facts agree with the theoretical analysis in Section 3, 4, and 5 very well.

Table 1. The CPU memory costs for different h and Q : unit GB.

h	Algorithm 1	Algorithm 2			Algorithm 3		
		$Q = 1$	$Q = 2$	$Q = 3$	$Q = 1$	$Q = 2$	$Q = 3$
$\frac{1}{4}$	1.271E-1	1.009E-1	1.195E-1	1.221E-1	1.967E-3	1.007E-3	2.006E-3
$\frac{1}{8}$	1.655E0	1.277E0	1.676E0	1.685E0	8.136E-3	2.723E-3	5.859E-3
$\frac{1}{16}$	1.807E1	1.792E1	2.481E1	2.487E1	4.856E-2	8.097E-2	7.375E-2
$\frac{1}{32}$	out of memory	2.677E2	out of memory	out of memory	7.546E-1	1.161E0	1.626E0

In order to observe the storage changes of the three proposed algorithms with respect to the sample number M , we fix the FEM parameter $h = 1/16$, the stochastic perturbation parameter $\varepsilon = 10^{-4}$ and the ADMM iteration number $t = 20$. The CPU memory costs with different cases are presented in Table 2. From Table 2, we can see that (1) The memory costs of Algorithm 1 and 2 increase linearly with the sample number M , which roughly equals $\mathcal{O}((M+1)N^2)$. (2) The storage cost of Algorithm 3 basically remain unchanged with the increasing of the sample number, that is approximately $\mathcal{O}(N^2)$. Based on all the results exhibited in Table 1 and Table 2, we can conclude that Algorithm 3 is an efficient method for the optimal control problem with random coefficient.

Table 2. The CPU memory costs for different M and Q : unit GB.

M	Algorithm 1	Algorithm 2			Algorithm 3		
		$Q = 1$	$Q = 2$	$Q = 3$	$Q = 1$	$Q = 2$	$Q = 3$
1	6.143E-2	5.456E-2	8.195E-2	7.839E-2	4.514E-2	6.814E-2	5.881E-2
10	2.498E-1	2.511E-1	2.964E-1	2.976E-1	4.568E-2	6.785E-2	5.901E-2
10^2	1.927E0	1.929E0	2.511E0	2.513E0	4.603E-2	6.887E-2	6.074E-2
10^3	1.807E1	1.792E1	2.481E1	2.487E1	4.856E-2	6.931E-2	7.375E-2

7. Conclusions. In this paper, three algorithms are proposed for the elliptic equation constraint optimal control problem with random coefficient. The Algorithm 1 is based on the FEM discretization, the MC approximation, and the ADMM iteration, which has global convergence. However, the computational cost and the storage cost are both very high. As an improvement, the Algorithm 2 is proposed, which reduce the computational cost by using the MME technique, but memory cost remain high. To overcome the disadvantage of Algorithm 2, Algorithm 3 is proposed, which reduce the memory cost significantly by transforming the constraint with random term to a deterministic one. Numerical simulations are carried out to verify the theoretical analysis, which illustrate the efficiency of the Algorithm 3 further. The applications of the proposed algorithms to other random optimal control problems will be discussed in our future works.

Acknowledgments. The work of H. Song was supported by the NSF of China under the grant No.11701210, the science and technology department project of Jilin Province under the grant No. 20190103029JH and 20200201269JC. the education department project of Jilin Province under the grant No. JJKH20180113KJ, and the fundamental research funds for the Central Universities. The work of X. Pang was supported by the NSF of China under the grant No. 11871245, 11771179, and 11726102. The work of J. Zhang was supported by the NSF of China under the grant No. 11971221 and 11731006, the Shenzhen Sci-Tech Fund under the grant No. JCYJ20170818153840322 and JCYJ20190809150413261, and Guangdong Provincial Key Laboratory of Computational Science and Material Design under the grant No. 2019B030301001. The authors also wish to thank the High Performance Computing Center of Jilin University, the Computing Center of Jilin Province, and the Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education for essential computing support.

Appendix. The proof of Theorem 7.

Proof. Using the definitions of the $F(y, u)$ and $F_{M,h}(y_h, u_h)$, and subtracting (19) from (1), we can get

$$\begin{aligned} & \mathbb{E} [|F(y^*, u^*) - F_{M,h}(y_{M,h}^*, u_{M,h}^*)|] \\ & \leq \mathbb{E} \left[\left| \mathbb{E} \left[\frac{1}{2} \int_D |y^* - y_d|^2 dx \right] - \mathbb{E}_M \left[\frac{1}{2} \int_D |y_{M,h}^* - y_d|^2 dx \right] \right| \right] \\ & \quad + \left| \frac{\gamma}{2} \int_D |u^* - u_d|^2 dx - \frac{\gamma}{2} \int_D |u_{M,h}^* - u_d|^2 dx \right| = \frac{1}{2}(I + II). \end{aligned} \quad (50)$$

For the first term i , we have

$$\begin{aligned} I &= \mathbb{E} [| \mathbb{E} [\|y^* - y_d\|_{L^2(D)}^2] - \mathbb{E}_M [\|y_{M,h}^* - y_d\|_{L^2(D)}^2] |] \\ & \leq \mathbb{E} [| (\mathbb{E} - \mathbb{E}_M) [\|Tu^* - y_d\|_{L^2(D)}^2] |] \\ & \quad + \mathbb{E} [| \mathbb{E}_M [\|Tu^* - y_d\|_{L^2(D)}^2] - \mathbb{E}_M [\|T_h u_{M,h}^* - y_d\|_{L^2(D)}^2] |] \\ & = I_1 + I_2. \end{aligned}$$

By using the MC error estimate (18), the triangle inequality, the estimate (7) with $j = 4$, and the inequality (52), it yields

$$\begin{aligned} I_1 &= \mathbb{E} [| (\mathbb{E} - \mathbb{E}_M) [\|Tu^* - y_d\|_{L^2(D)}^2] |] \\ & \leq \left(\mathbb{E} [((\mathbb{E} - \mathbb{E}_M) [\|Tu^* - y_d\|_{L^2(D)}^2])^2] \right)^{\frac{1}{2}} \\ & \leq \left(\mathbb{E} [(\mathbb{E} - \mathbb{E}_M) [\|Tu^* - y_d\|_{L^2(D)}^4]] \right)^{\frac{1}{2}} \\ & \leq \frac{1}{\sqrt{M}} \left(\mathbb{E} [\|Tu^* - y_d\|_{L^2(D)}^4] \right)^{\frac{1}{4} \cdot 2} \\ & \leq \frac{1}{\sqrt{M}} \left(\mathbb{E} [(\|Tu^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})^4] \right)^{\frac{1}{4} \cdot 2} \\ & \leq \frac{1}{\sqrt{M}} \left((\mathbb{E} [8 \|Tu^*\|_{L^2(D)}^4])^{\frac{1}{4}} + (\mathbb{E} [8 \|y_d\|_{L^2(D)}^4])^{\frac{1}{4}} \right)^2 \\ & \leq \frac{1}{\sqrt{M}} \sqrt{8} (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})^2. \end{aligned} \quad (51)$$

Applying the Cauchy-Schwarz inequality, the triangle inequality, the MC error estimate (18), and the inequality

$$(\mathbb{E}[(m+n)^4])^{\frac{1}{4}} \leq (\mathbb{E}[8(m^4+n^4)])^{\frac{1}{4}} \leq (\mathbb{E}[8m^4])^{\frac{1}{4}} + (\mathbb{E}[8n^4])^{\frac{1}{4}}, \quad \forall m, n \geq 0, \quad (52)$$

we can get

$$\begin{aligned} I_2 &= \mathbb{E} [| \mathbb{E}_M [-\|T_h u_{M,h}^* - Tu^*\|_{L^2(D)}^2 + 2(Tu^* - T_h u_{M,h}^*, Tu^* - y_d)] |] \\ & \leq 2 \mathbb{E} [\mathbb{E}_M [\|Tu^* - T_h u_{M,h}^*\|_{L^2(D)} \|Tu^* - y_d\|_{L^2(D)}]] \\ & \quad + \mathbb{E} [\mathbb{E}_M [\|T_h u_{M,h}^* - Tu^*\|_{L^2(D)}^2]] \\ & = I_{21} + I_{22}, \end{aligned}$$

and

$$\begin{aligned}
 I_{21} &\leq 2\mathbb{E} [(\mathbb{E}_M - \mathbb{E}) [\|Tu^* - T_h u_{M,h}^*\|_{L^2(D)} (\|Tu^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})]] \\
 &\quad + 2\mathbb{E} [\|Tu^* - T_h u_{M,h}^*\|_{L^2(D)} (\|Tu^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})] \\
 &\leq \frac{2}{\sqrt{M}} \left(\mathbb{E} [(\|Tu^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})^2 \|Tu^* - T_h u_{M,h}^*\|_{L^2(D)}^2] \right)^{\frac{1}{2}} \\
 &\quad + 2 (\mathbb{E} [\|Tu^* - T_h u^*\|_{L^2(D)} \|Tu^*\|_{L^2(D)}] + \mathbb{E} [\|Tu^* - T_h u^*\|_{L^2(D)} \|y_d\|_{L^2(D)}]) \\
 &\quad + 2 (\mathbb{E} [\|T_h(u^* - u_{M,h}^*)\|_{L^2(D)} \|Tu^*\|_{L^2(D)}] + \mathbb{E} [\|T_h(u^* - u_{M,h}^*)\|_{L^2(D)} \|y_d\|_{L^2(D)}]) \\
 &\leq \frac{2}{\sqrt{M}} (\mathbb{E} [(\|Tu^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})^4])^{\frac{1}{4}} (\mathbb{E} [\|Tu^* - T_h u_{M,h}^*\|_{L^2(D)}^4])^{\frac{1}{4}} \\
 &\quad + 2 (\mathbb{E} [\|Tu^*\|_{L^2(D)}^2])^{\frac{1}{2}} (\mathbb{E} [\|(T - T_h)u^*\|_{L^2(D)}^2])^{\frac{1}{2}} \\
 &\quad + 2 (\mathbb{E} [\|y_d\|_{L^2(D)}^2])^{\frac{1}{2}} (\mathbb{E} [\|(T - T_h)u^*\|_{L^2(D)}^2])^{\frac{1}{2}} \\
 &\quad + 2 (\mathbb{E} [\|Tu^*\|_{L^2(D)}^2])^{\frac{1}{2}} (\mathbb{E} [\|T_h(u^* - u_{M,h}^*)\|_{L^2(D)}^2])^{\frac{1}{2}} \\
 &\quad + 2 (\mathbb{E} [\|y_d\|_{L^2(D)}^2])^{\frac{1}{2}} (\mathbb{E} [\|T_h(u^* - u_{M,h}^*)\|_{L^2(D)}^2])^{\frac{1}{2}}.
 \end{aligned}$$

Taking $j = 2, 4$ in (7), (14), and (15), it follows from (52) that

$$\begin{aligned}
 I_{21} &\leq \frac{2}{\sqrt{M}} \left((\mathbb{E} [8\|Tu^*\|_{L^2(D)}^4])^{\frac{1}{4}} + (\mathbb{E} [8\|y_d\|_{L^2(D)}^4])^{\frac{1}{4}} \right) \\
 &\quad \cdot \left((\mathbb{E} [8\|(T - T_h)u^*\|_{L^2(D)}^4])^{\frac{1}{4}} + (\mathbb{E} [8\|T_h(u^* - u_{M,h}^*)\|_{L^2(D)}^4])^{\frac{1}{4}} \right) \\
 &\quad + 2 (C_1 C_3 h^2 \|u^*\|_{L^2(D)}^2 + \|y_d\|_{L^2(D)} C_3 h^2 \|u^*\|_{L^2(D)}) \\
 &\quad + 2 (C_1 C_2 \|u^* - u_{M,h}^*\|_{L^2(D)} + C_2 \|y_d\|_{L^2(D)} \|u^* - u_{M,h}^*\|_{L^2(D)}) \\
 &= \frac{2}{\sqrt{M}} \left(\sqrt[4]{8} (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)}) \right) \\
 &\quad \cdot \left(\sqrt[4]{8} (C_3 h^2 \|u^*\|_{L^2(D)} + C_2 \|u^* - u_{M,h}^*\|_{L^2(D)}) \right) \\
 &\quad + 2 ((C_1 C_3 \|u^*\|_{L^2(D)}^2 + C_3 \|y_d\|_{L^2(D)} \|u^*\|_{L^2(D)}) h^2) \\
 &\quad + 2 ((C_1 C_2 \|u^*\|_{L^2(D)} + C_2 \|y_d\|_{L^2(D)}) \|u^* - u_{M,h}^*\|_{L^2(D)}).
 \end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
 I_2 &\leq 2 (C_1 C_3 \|u^*\|_{L^2(D)}^2 + C_3 \|y_d\|_{L^2(D)} \|u^*\|_{L^2(D)} + (C_1 C_2 \|u^*\|_{L^2(D)} + C_2 \|y_d\|_{L^2(D)}) K_2) h^2 \\
 &\quad + \frac{4\sqrt{2}}{\sqrt{M}} ((C_1 C_3 \|u^*\|_{L^2(D)} + C_3 \|y_d\|_{L^2(D)} + C_1 C_2 K_2) \|u^*\|_{L^2(D)} + C_2 K_2 \|y_d\|_{L^2(D)}) h^2 \\
 &\quad + \frac{4\sqrt{2}}{\sqrt{M}} (C_2 K_1 (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})) \\
 &\quad + \frac{2}{\sqrt{M}} (K_1 C_1 C_2 \|u^*\|_{L^2(D)} + K_1 C_2 \|y_d\|_{L^2(D)}) \\
 &\quad + \frac{4\sqrt{2}}{\sqrt{M}} h^4 (C_3^2 \|u^*\|^4 + C_2^2 K_2^2) + \frac{4\sqrt{2}}{\sqrt{M}} \frac{1}{M} C_2^2 K_1^2 + \frac{4\sqrt{2}}{M} C_2^2 K_1^2 \\
 &\quad + 4\sqrt{2} h^4 (C_3^2 \|u^*\|^4 + C_2^2 K_2^2).
 \end{aligned}$$

When $M \gg 1$ and $h \ll 1$, we have

$$\begin{aligned}
 I_2 \leq & 2(C_1 C_3 \|u^*\|_{L^2(D)}^2 + C_3 \|y_d\|_{L^2(D)} \|u^*\|_{L^2(D)} \\
 & + (C_1 C_2 \|u^*\|_{L^2(D)} + C_2 \|y_d\|_{L^2(D)}) K_2) h^2 \\
 & + \frac{4\sqrt{2}}{\sqrt{M}} ((C_1 C_3 \|u^*\|_{L^2(D)} + C_3 \|y_d\|_{L^2(D)} + C_1 C_2 K_2) \|u^*\|_{L^2(D)} + C_2 K_2 \|y_d\|_{L^2(D)}) \\
 & + \frac{4\sqrt{2}}{\sqrt{M}} (C_2 K_1 (C_1 \|u^*\|_{L^2(D)} + \|y_d\|_{L^2(D)})) \\
 & + \frac{2}{\sqrt{M}} (K_1 C_1 C_2 \|u^*\|_{L^2(D)} + K_1 C_2 \|y_d\|_{L^2(D)}) \\
 & + \frac{4\sqrt{2}}{\sqrt{M}} (C_3^2 \|u^*\|^4 + C_2^2 K_2^2 + 2C_2^2 K_1^2) + 4\sqrt{2} h^2 (C_3^2 \|u^*\|^4 + C_2^2 K_2^2).
 \end{aligned}$$

Next, we deal with the second term, applying the Cauchy-Schwarz inequality, triangle inequality, and the estimate (21), we obtain

$$\begin{aligned}
 II = & \gamma |(u^* - u_d, u^* - u_d)_{L^2(D)} - (u_{M,h}^* - u_d, u_{M,h}^* - u_d)_{L^2(D)}| \\
 = & \gamma | -\|u_{M,h}^* - u^*\|_{L^2(D)}^2 + 2(u^* - u_{M,h}^*, u^* - u_d)_{L^2(D)} | \\
 \leq & 2\gamma \|u^* - u_{M,h}^*\|_{L^2(D)} \|u^* - u_d\|_{L^2(D)} \\
 \leq & 2\gamma (\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)}) \|u^* - u_{M,h}^*\|_{L^2(D)} \\
 \leq & 2\gamma (\|u^*\|_{L^2(D)} + \|u_d\|_{L^2(D)}) (K_1 \frac{1}{\sqrt{M}} + K_2 h^2).
 \end{aligned}$$

Substituting I_1 , I_2 , and II into (50), the conclusion is confirmed naturally. \square

REFERENCES

- [1] I. Babuška, F. Nobile and R. Tempone, [A stochastic collocation method for elliptic partial differential equations with random input data](#), *SIAM Rev.*, **52** (2010), 317–355.
- [2] I. Babuška, F. Nobile and R. Tempone, [A stochastic collocation method for elliptic partial differential equations with random input data](#), *SIAM J. Numer. Anal.*, **45** (2007), 1005–1034.
- [3] I. Babuška, R. Tempone and G. E. Zouraris, [Solving elliptic boundary value problems with uncertain coefficients by the finite element method: The stochastic formulation](#), *Comput. Methods. Appl. Mech. Engrg.*, **194** (2005), 1251–1294.
- [4] A. Barth, C. Schwab and N. Zollinger, [Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients](#), *Numer. Math.*, **119** (2011), 123–161.
- [5] R. E. Caflisch, [Monte-Carlo and quasi-Monte Carlo methods](#), in *Acta Numerica*, Acta Numer., 7, Cambridge University Press, Cambridge, 1998, 1–49.
- [6] X. Cai and D. Han, [\$O\(1/t\)\$ complexity analysis of the alternating direction method of multipliers](#), *Sci. China Math.*, **62** (2019), 795–808.
- [7] Y. Cao, M. Y. Hussaini and T. A. Zang, [An efficient Monte Carlo method for optimal control problems with uncertainty](#), *Comput. Optim. Appl.*, **26** (2003), 219–230.
- [8] P. Chen, U. Villa and O. Ghattas, [Taylor approximation and variance reduction for PDE-constrained optimal control under uncertainty](#), *J. Comput. Phys.*, **385** (2019), 163–186.
- [9] P. G. Ciarlet, [The Finite Element Method for Elliptic Problems](#), Studies in Mathematics and its Applications, 4, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.
- [10] X. Feng, J. Lin and C. Lorton, [An efficient numerical method for acoustic wave scattering in random media](#), *SIAM/ASA J. Uncertain. Quantif.*, **3** (2015), 790–822.
- [11] X. Feng, J. Lin and C. Lorton, [A multimodes Monte Carlo finite element method for elliptic partial differential equations with random coefficients](#), *Int. J. Uncertain. Quantif.*, **6** (2016), 429–443.
- [12] D. Gilbarg and N. S. Trudinger, [Elliptic Partial Differential Equations of Second Order](#), Classics in Mathematics, Springer-Verlag, Berlin, 2001.
- [13] M. D. Gunzburger, H.-C. Lee and J. Lee, [Error estimates of stochastic optimal Neumann boundary control problems](#), *SIAM J. Numer. Anal.*, **49** (2011), 1532–1552.

- [14] Y. Hao, H. Song, X. Wang and K. Zhang, [An alternating direction method of multipliers for the optimization problem constrained with a stationary Maxwell system](#), *Commun. Comput. Phys.*, **24** (2018), 1435–1454.
- [15] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich, *Optimization with PDE Constraints*, Mathematical Modelling: Theory and Applications, 23, Springer, New York, 2009.
- [16] L. S. Hou, J. Lee and H. Manouzi, [Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs](#), *J. Math. Anal. Appl.*, **384** (2011), 87–103.
- [17] D. P. Kouri, M. Heinkenschloss, D. Ridzal and B. G. van Bloemen Waanders, [A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty](#), *SIAM J. Sci. Comput.*, **35** (2013), A1847–A1879.
- [18] A. Kunoth and C. Schwab, [Analytic regularity and GPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs](#), *SIAM J. Control Optim.*, **51** (2013), 2442–2471.
- [19] J. Li, X. Wang and K. Zhang, [An efficient alternating direction method of multipliers for optimal control problems constrained by random Helmholtz equations](#), *Numer. Algorithms*, **78** (2018), 161–191.
- [20] P. Li and S. T. Yau, [On the Schrödinger equation and the eigenvalue problem](#), *Comm. Math. Phys.*, **88** (1983), 309–318.
- [21] J. C. De los Reyes, *Numerical PDE-Constrained Optimization*, SpringerBriefs in Optimization, Springer, Cham, 2015.
- [22] L. J. Roman and M. Sarkis, [Stochastic Galerkin method for elliptic SPDEs: A white noise approach](#), *Discrete Contin. Dyn. Syst. Ser. B*, **6** (2006), 941–955.
- [23] W. Shen, T. Sun, B. Gong and W. Liu, [Stochastic Galerkin method for constrained optimal control problem governed by an elliptic integro-differential PDE with stochastic coefficients](#), *Int. J. Numer. Anal. Model.*, **12** (2015), 593–616.
- [24] H. F. Weinberger, [Lower bounds for higher eigenvalues by finite difference methods](#), *Pacific J. Math.*, **8** (1958), 339–368.
- [25] D. Xiu, [Fast numerical methods for robust optimal design](#), *Eng. Optim.*, **40** (2008), 489–504.
- [26] K. Zhang, J. Li, Y. Song and X. Wang, [An alternating direction method of multipliers for elliptic equation constrained optimization problem](#), *Sci. China Math.*, **60** (2017), 361–378.

Received March 2020; 1st revision March 2020; 2nd revision April 2020.

E-mail address: pangxw16@mails.jlu.edu.cn

E-mail address: songhaiming@jlu.edu.cn

E-mail address: xxwang@ualr.edu

E-mail address: zhangjc@sustech.edu.cn