INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 12, Number 4, Pages 593–616

STOCHASTIC GALERKIN METHOD FOR CONSTRAINED OPTIMAL CONTROL PROBLEM GOVERNED BY AN ELLIPTIC INTEGRO-DIFFERENTIAL PDE WITH RANDOM COEFFICIENTS

W. SHEN, T. SUN, B. GONG, AND WENBIN LIU

Abstract. In this paper, a stochastic finite element approximation scheme is developed for an optimal control problem governed by an elliptic integro-differential equation with random coefficients. Different from the well-studied optimal control problems governed by stochastic PDEs, our control problem has the control constraints of obstacle type, which is mostly seen in real applications. We develop the weak formulation for this control and its stochastic finite element approximation scheme. We then obtain necessary and sufficient optimality conditions for the optimal control and the state, which are the base for deriving a priori error estimates of the approximation in our work. Instead of using the infinite dimensional Lagrange multiplier theory, which is currently used in the literature but often difficult to handle inequality control constraints, we use a direct approach by applying the well-known Lions' Lemma to the reduced optimal problem. This approach is shown to be applicable for a wide range of control constraints. Finally numerical examples are presented to illustrate our theoretical results.

Key words. Priori error estimates, stochastic Galerkin method, optimal control problem, integro-differential equation, constraint of obstacle type.

1. Introduction

Optimal control problems governed by partial different equations have been a major research topic in applied mathematics and control theory. Since the milestone work of J.P Lions [33], a great deal of progress has been made in many aspects such as stability, observability and numerical methods, which are too extensive to be mentioned here even very briefly. Finite element approximation of optimal control problems plays a very important role in numerical methods for these problems, and, the finite element approximation of optimal control problems governed by various partial differential equations, either linear or nonlinear, have been much studied in the literature. For optimal control problems governed by the classic PDEs, the optimality conditions and their finite element approximation and a prior error estimates were established long ago, for example, see the early work in [11]. There have been extensive studies on this aspect for such as elliptic equations, parabolic equations, Stokes equations, and Niavoir-Stokes equations. Some of recent progress in this area has been summarized in [20, 27, 31, 35, 43, 46, 56], and the references cited therein. Systematic introductions of the finite element method for PDEs and optimal control problems can be found in, for example, [43, 46, 56]. There also exists an extensive body of studies adaptive finite element methods for various optimal control problems, which is again too extensive to be mentioned here even very briefly. For a recent summary in computational optimal control, we refer our readers to the recent monograph [36].

Received by the editors December 2, 2014.
2000 *Mathematics Subject Classification.* 65K15, 49M05.
This work is supported by the NSF of China (No. 11271231, 11301300 and 11326226). The first author is partially supported by the China Scholarship Council.

Recently, optimal control problems with more complicated state equations have been considered, particularly those with the integro-differential state equations. Integro-differential equations and their control of this nature appear in applications such as heat conduction in materials with memory, population dynamics, and viscous-elasticity; cf., e.g., Friedman and Shinbrot [12], Heard [21], and Renardy, Hrusa, and Nohel [47]. For equations with nonsmooth kernels, we refer to Grimmer and Pritchard [17], Lunardi and Sinestrari [39], and Lorenzi and Sinestrari [38] and references therein. One very important characteristic of all these models is that they all express a conservation of a certain quantity mass momentum in any moment for any subdomain. This in many applications is the most desirable feature of the approximation method when it comes to numerical solution of the corresponding initial boundary value problem. Furthermore finite element methods for parabolic integro-differential equations problems with a smooth kernel have been discussed in, e.g., Cannon and Lin [6], LeRoux and Thomée [30], Lin, Thomée, and Wahlbin [32], Sloan and Thomée [54], Thomée and Zhang [55], and Yanik and Fairweather [61].

Only very recently the finite element approximation of optimal control with the integro-differential state equations has been systematically studied. For example, the finite element method for the optimal control governed by elliptic integral equations and integro-differential equations has been made in [22], in which the a priori and a posteriori error estimations were obtained. For optimal control problems governed by linear parabolic (and quasi-parabolic) integro-differential equations, a priori error estimates of finite element approximation were studied in [51, 52], hyperbolic integro-differential equations [53]. It is, however, much more difficult to study adaptive finite element methods for control problems governed by linear parabolic integro-differential equations.

Uncertainty, such as uncertain parameters, arises in many complex real-world problems of physical and engineering interests. It is well known that these problems can be described by different kinds of stochastic partial differential equations (SPDEs). In recent years, finite element methods for stochastic elliptic and parabolic PDEs (here we mean the equations with stochastic perturbation in their coefficients.) have been a subject of growing interest in the scientific community (see e.g. [1, 2, 8, 50]), which have been widely used to model fluid flows in porous media in many areas, e.g., transport of pollutants in groundwater and oil recovery processes.

The well known Monte Carlo (MC) method is still the most popular method for simulating stochastic elliptic PDEs and dealing with the statistic characteristics of the solution, although it is a rather computationally expensive method (see e.g. [9, 45]) for higher accuracy. Other alternatives to Monte Carlo method have been employed in the field of stochastic mechanics. A popular technique is the perturbation method, cf. [26]. Given certain smoothness conditions, the random functions and operators involved in the differential equation are expanded in a Taylor series about their respective mean values. Another approach is the Neumann expansion series method, e.g. [1]. In this method the inverse of the boundary value problems stochastic operator is approximated by its Neumann series. Based on a spectral representation of the uncertainty, the spectral stochastic finite element method (SSFEM), e.g. [16] was introduced. This method utilized the Karhunen-Loève expansion of correlated random functions, (cf. [37]), and obtaind the solution by a

Galerkin method in a space of stochastic functions. More information, references and reviews on stochastic finite elements can be found in [41]-[42].

Following [16], the stochastic Galerkin method has been applied to various stochastic problems, e.g. [3, 14, 25, 50, 59, 60]. Such a Galerkin method allows us to use polynomial chaos (PC) or generalized polynomial chaos (gPC) serving as a complete basis to represent random processes as explicit functional of finite number of independent random variables. Then, an explicit functional relationship between the independent random variables and the solution is achieved and the stochastic governing equations are transformed to a set of deterministic equations which can be readily discretized via more standard numerical techniques like the finite element or adaptive finite element methods for higher accuracy.

Here we utilize the stochastic Galerkin method to approximate the optimal control governed by the PDE with random fields in its coefficients. In many applications,optimization of physical and engineering systems can be formulated as optimal control problems that are constrained by such stochastic PDEs. Nevertheless, the development of stochastic optimal control problem constrained by the stochastic PDEs can still be considered to be in its infancy, see the very recent work of [19]-[29]. The work [19] dealed with deterministic Neumann boundary control in order to obtain a priori error estimate for the numerical approximation of stochastic steady diffusion problem. In [48], numerical experiments were conducted with 'pure' stochastic control function as well as 'semi' stochastic control function for an optimal control problem constrained by stochastic steady diffusion problem. In [23] and [29], stochastic optimal control problems constrained by stochastic elliptic PDEs with deterministic distributed control function are introduced. The authors prove the existence of the optimal solution, establish the validity of the Lagrange multiplier rule and obtain stochastic optimality system. Then, they use the Wiener-Itô (W-I) chaos or the Karhunen-Loève (K-L) expansion as a main tool to convert stochastic optimality system to deterministic optimality system. Finally, a priori error estimates for Galerkin approximation of the optimality system in both physical space and stochastic space are provided. Let us emphasize these works use the infinite dimensional Lagrange multiplier theory to deal with the PDE constraint in the control problems, which is often complicated to extend to deal with the cases where the control has inequality constraints such as obstacle types.

In this paper, we deal with a priori error estimate of Stochastic Galerkin Method for optimal control problem governed by an integro-differential equation with random fields in its coefficients, and with the control constraints of obstacle type, which is mostly seen in applications. The objective is of a state tracking type adding the weighted energy of tracking, and the deterministic control is of the obstacle constrained type. In Section 2, we introduce some function spaces and the stochastic optimal control problem. In Section 3, we represent the stochastic elliptic PDE in term of the generalized polynomial chaos (gPC) expansion and obtain the finite dimensional optimal control problem. The important feature of this work is that: instead of using the Lagrange multiplier method in infinite dimensional spaces, which is currently used in the literature but often difficult to deal with inequality control constraints, we use a direct approach by applying the well-known Lions' Lemma to the reduced optimal problem. We then obtain the necessary and sufficient optimality conditions. After constructing finite element spaces and theirs approximation properties with respect to both the spatial space and the probability space, the stochastic Galerkin approximation scheme is established in Section 4. Section 5 considers a priori error estimates for the state, the co-state and the control variables. Numerical examples are presented to illustrate our theoretical results in Section 6.

2. Model control problem

2.1. Function spaces and notations. Let $D \subset \mathbb{R}^d$ be a convex bounded polygonal spatial domain with boundary ∂D and $1 \leqslant d \leqslant 3$. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space, where Ω is a set of outcomes, $\mathcal{F}$ is a σ-algebra of events and $\mathcal{P} : \mathcal{F} \to [0,1]$ is a probability measure. Denote by $B(D)$ the Borel σ-algebra generated by the open subset of D.

Throughout the paper, we adopt the standard notations for Sobolev spaces as in [28, 33, 34, 57], such as $W^{m,q}(\Omega)$ on Ω with norm $\|\cdot\|_{m,q,\Omega}$, and semi-norm$|\cdot|_{m,q,\Omega}$ for $1 \leqslant q \leqslant \infty$. Set $W_0^{m,q}(\Omega) = \Big\{w \in W^{m,q}(\Omega) : w|_{\partial\Omega} = 0\Big\}$. Also denote $W^{m,2}(\Omega)(W_0^{m,2}(\Omega))$ by $H^m(\Omega)$ $(H_0^m(\Omega))$, with norm$\|\cdot\|_{m,\Omega}$, and semi-norm $|\cdot|_{m,\Omega}$. Denote by $L^s(0,T;W^{m,q}(\Omega))$ the Banach space of all L^s integrable functions from $(0,T)$ into $W^{m,q}(\Omega)$ with norm $\| \ v \ \|_{L^s(0,T;W^{m,q}(\Omega))}= \ (\int_0^T \| \ v \ \|^s_{W^{m,q}(\Omega)} \ dt)^{\frac{1}{s}}$ for $s \in [1,\infty)$ and the standard modification for $s = \infty$. Similarly, one can define the spaces $H^1(0,T;W^{m,q}(\Omega))$ and $C^k(0,T;W^{m,q}(\Omega))$. The details can be found in [34].

With these standard Sobolev spaces, throughout this paper, we adopt the definition of stochastic Sobolev spaces (see [3, 23, 29]). For a nonnegative integer s and $1 \leqslant q < +\infty$, let $L^q(\Omega; W^{s,q}(D))$ contain stochastic functions, $v : \ \Omega \times D \to R$, that are measurable with respect to the product σ-algebra $\mathcal{F} \otimes B(D)$ and equipped with the averaged norms

$$\|v\|_{L^q(\Omega;W^{s,q}(D))} = \Big(\mathbb{E}[\|v\|^q_{W^{s,q}(D)}]\Big)^{1/q} = \Big(\mathbb{E}[\sum_{|\alpha|\leqslant s}\int_D |\partial^\alpha v|^q dx]\Big)^{1/q},$$

and

$$\|v\|_{L^\infty\Omega;W^{s,\infty}(D))} = \max_{|\alpha|\leqslant s}\big(\text{ess}\sup_{\Omega\times D} |\partial^\alpha v|\big),$$

where $\mathbb{E}$ is the expected value. Observe that if $v \in L^q(\Omega; W^{s,q}(D))$, then $v(\cdot,\omega) \in W^{s,q}(D)$ almost surly (a.s.) and $\partial^\alpha v(x,\cdot) \in L^q(\Omega)$ a.e. on D for $|\alpha| \leqslant s$.

When $s = 0$, $q = 2$, the above space is just

$$L^2(\Omega; L^2(D)) = \{v : D \times \Omega \to \mathbb{R} \mid \|v\|_{L^2(\Omega;L^2(D))} < \infty\},$$

with the norm

$$\|v\|^2_{L^2(\Omega;L^2(D))} = \int_\Omega \|v\|^2_{L^2(D)} d\mathcal{P} = \mathbb{E}\|v\|^2_{L^2(D)}.$$

One can define spaces $L^2(\Omega; H^1(D))$ and $L^2(\Omega; H_0^1(D))$ similarly. Note that these stochastic Sobolev spaces are Hilbert spaces.

2.2. Model stochastic optimal control problem. In our work, the following control problem governed by an elliptic integro-differential equation with random coefficients and a control constraint of an obstacle type:

$$\min_{u\in K} \mathcal{J}(u) = \min_{u\in K} \mathbb{E}\big(\frac{1}{2}\int_D |y - y_d|^2 dx + \frac{\alpha}{2}\int_D |u|^2 dx\big) \tag{1}$$

subject to

(2)

$$\begin{cases} -\nabla\cdot[a(x,\omega)\nabla y(x,\omega)] - \int_D g(s,x,\omega)y(s,\omega)ds = u(x), & x \in D, \ \omega \in \Omega, \\ y(x,\omega) = 0, & x \in \partial D, \ \omega \in \Omega, \end{cases}$$

where $\mathcal{J}$ is a cost functional, $y : \bar{D} \times \Omega \to \mathbb{R}$ is the state variable, $y_d : \bar{D} \times \Omega \to \mathbb{R}$ is a given target solution, $a : D \times \Omega \to \mathbb{R}$ is a stochastic function with continuous and bounded covariance function, $g(\cdot,\cdot,\cdot)\ D \times D \times \Omega \to \mathbb{R}$ is another stochastic function with continuous and bounded covariance function (precise regularity conditions will be given later), $u : D \to \mathbb{R}$ is a deterministic control, α is a positive constant measuring the importance between two terms in $\mathcal{J}$. The operator ∇ means the gradient with respect to the spatial variable $x \in D$ only. Here, K is a closed convex subset in the control space $L^2(D)$. In the following context, we will discuss some different cases on the choice of K.

If we denote by $B(D)$ the Borel σ-algebra generated by the open subsets of D, then a and g are assumed measurable with respect to the σ-algebras $(B(D) \otimes \mathcal{F})$ and $(B(D) \otimes B(D) \otimes \mathcal{F})$, and independent. To ensure well-postness of the state equation, we assume that there are positive constants a_{min} and a_{max}, such that

$$a_{min} \leqslant a(x,\omega) \leqslant a_{max}, \quad \text{a. e. } (x,\omega) \in D \times \Omega. \tag{3}$$

We also assume that there exists a positive constant $c_0 > 0$ such that

$$|g(s,x,\omega)| \leqslant c_0, \quad \text{a. e. } (s,x,\omega) \in D \times D \times \Omega, \tag{4}$$

where c_0 is such that there is a positive constant ε satisfying

$$\int_D a\nabla v \nabla v \geqslant (c_0+\varepsilon)|D| \parallel v \parallel_{1,D}^2, \ \forall\, v \in H^1(D),\ \varepsilon > 0, \tag{5}$$

where $|D| = \int_D 1$. Then with these assumptions, the existence and uniqueness of a weak solution y for (2) in a sense of average can be set and proved as below.

In the following, we will take the state space $Y = L^2(\Omega; H_0^1(D))$ and the control space $U = L^2(D)$. In addition, C will denote general constants. To present the weak formulation for equation (2) after average, we introduce the following bilinear forms:

$$\begin{aligned} A[y,v] &= \mathbb{E}[\int_D a(x,\omega)\nabla y(x,\omega)\cdot\nabla v(x,\omega)dx] \\ &= \mathbb{E}[\int_D a\nabla y\cdot\nabla v dx], \quad \forall\ y,\ v \in Y, \\ G[y,v] &= \mathbb{E}[\int_D(\int_D g(s,x,\omega)y(s,\omega)v(x,\omega)ds)dx] \\ &= \mathbb{E}[\int_D(\int_D g(s,x)y(s)v(x)ds)dx], \quad \forall\ y,\ v \in L^2(\Omega; L^2(D)), \end{aligned} \tag{6}$$

and

$$\begin{aligned} [u,v] &= \mathbb{E}(\int_D u(x,\omega)v(x,\omega)dx) = \mathbb{E}(\int_D uv dx) \quad \forall\ u,\ v \in L^2(\Omega; L^2(D)), \\ (u,v) &= \int_D u(x)v(x)dx, \quad \forall\ u,\ v \in U. \end{aligned} \tag{7}$$

Then, we can derive the weak formulation of optimal control problem (1)-(2) by:

$$\min_{u\in K} \mathcal{J}(u) = \min_{u\in K} \mathbb{E}\big(\frac{1}{2}\int_D |y-y_d|^2 dx + \frac{\alpha}{2}\int_D |u|^2 dx\big) \tag{8}$$

subject to

$$A[y,v] - G[y,v] = [u,v], \quad \forall\ v \in Y. \tag{9}$$

Here from the Lax-Milgram lemma (see [5]), we have the following theorem about the existence and uniqueness of the solution for (9).

Theorem 2.1. *Let $u \in L^2(\Omega, L^2(D))$. Then there exists a unique solution for the following weak equations: find $y, q \in Y$, such that:*

$$A[y,v] - G[y,v] = [u,v], \quad \forall\ v \in Y, \tag{10}$$

and the dual equation:

$$A[q,p] - G[q,p] = [u,q], \quad \forall\ q \in Y. \tag{11}$$

Proof. Note that from ellipticity condition (3) on a and the assumption (4) on g , there exist a positive constant $C > 0$ such that

$$|A[y,v] - G[y,v]| \leqslant C \parallel y \parallel_{L^2(\Omega,H^1(D))} \parallel v \parallel_{L^2(\Omega,H^1(D))}, \quad \forall\ v \in Y. \tag{12}$$

Furthermore by applying the Schwarz inequality

$$G[y,y] \leqslant c_0 \parallel y \parallel^2_{L^2(\Omega,L^2(D))} \leqslant c_0|D| \parallel y \parallel^2_{L^2(\Omega,H^1(D))} \quad \forall\ y \in Y. \tag{13}$$

Thus it follows from the assumption (4) that there exists a positive constant $c > 0$ such that

$$A[v,v] - G[v,v] \geqslant c \parallel v \parallel^2_{L^2(\Omega,H^1(D))}, \quad \forall\ v \in Y. \tag{14}$$

On the other hand, we can easily see that there is a constant $C > 0$ such that

$$|[u,v]| \leqslant C \parallel u \parallel_{L^2(\Omega,L^2(D))} \parallel v \parallel_{L^2(\Omega,H^1(D))}, \quad \forall\ v \in Y. \tag{15}$$

Hence, by the Lax-Milgram lemma, (9) has a unique solution. The equation (11) can be dealt with similarly. □

Remark 2.2. *The assumption* (4) *can be further relaxed to (see* [22]*): There exists a constant $c > 0$ such that for any $v \in Y$ exists a $w \in Y$*

$$A[v,w] - G[v,w] \geqslant c \parallel v \parallel_{L^2(\Omega,H^1(D))} \parallel w \parallel_{L^2(\Omega,H^1(D))}. \tag{16}$$

Then (9) *has a unique solution. Furthermore the a priori error estimates derived in Section 5 still hold if a discreitizd version* (16) *holds . For the details the readers are referred to* [22].

2.3. Optimality Conditions. It is essential to derive the optimality conditions for the above constrained optimal control problem in order to set up its suitable finite element approximation and obtain error estimates. In [23, 29], stochastic control problems with un-constrained control were studied, by using the infinite dimensional Lagrange multiplier theory. It is not trivial to apply the infinite dimensional Lagrange multiplier theory to our cases. Here, we use a different approach, which is much simpler and widely used in the literature (see [36, 51], to be explained below:

It is well-known that the PDE-constrained optimal control problem (8)-(9) can be recast to be the following minimization problem:

$$\min_{u \in K} \mathcal{J}(u)$$

where

$$\mathcal{J}(u) = \mathbb{E}\Big(\frac{1}{2}\int_D |y(u) - y_d|^2 dx + \frac{\alpha}{2}\int_D |u|^2 dx\Big) \tag{17}$$

subject to

$$A[y(u),v] - G[y(u),v] = [u,v], \quad \forall\ v \in Y. \tag{18}$$

Here Y and U are Hilbert spaces, functional $\mathcal{J}$ is strictly convex and $K \subset U$ is a closed convex set. According to Lions' theorem ([33]), there exists a unique minimizer, which satisfies the following variational inequality:

$$\mathcal{J}'(u)(w - u) \geqslant 0, \quad \forall\ w \in K. \tag{19}$$

Here, the directional derivative of functional $\mathcal{J}$ at $u \in K$ along the direction $w \in K$ is defined by

$$\mathcal{J}'(u)(w) = \lim_{t \to 0^+} \frac{\mathcal{J}'(u+tw) - \mathcal{J}'(u)}{t}. \tag{20}$$

Applying the above theory to our control problem, we have the following theorem.

Theorem 2.3. *There exists a unique solution of* (8)-(9). *Furthermore a pair* (y, u) *is the solution iff there is a co-state variable* $p \in Y$*, such that the triplet* (y, p, u) *satisfies the following optimality system:*

$$\begin{cases} A[y, v] - G[y, v] = [u, v], \quad \forall\, v \in Y, \\ A[q, p] - G[q, p] = [y - y_d, q], \quad \forall\, q \in Y, \\ [p + \alpha u, w - u] \geqslant 0, \quad \forall\, w \in K. \end{cases} \tag{21}$$

Proof. Let $\mathcal{J}(u) = g(y(u)) + j(u)$, where $g(y(u)) = \mathbb{E}\big(\frac{1}{2}\int_D |y(u) - y_d|^2 dx\big)$ and $j(u) = \mathbb{E}(\frac{\alpha}{2}\int_D |u|^2 dx)$. Applying the variational inequality, the optimal condition reads

$$j'(u)(v-u) + (g(y(u)))'(v-u) \geqslant 0, \quad \forall\, v \in K. \tag{22}$$

It is clear that

$$\begin{aligned} j'(u)(v-u) &= \lim_{t \to 0^+} \frac{1}{t} E\Big(\frac{\alpha}{2}\int_D \big[|u + t(v-u)|^2 - |u|^2\big] dx\Big) \\ &= E\Big(\int_D \alpha u \cdot (v-u) dx\Big) = [\alpha u, v-u], \end{aligned} \tag{23}$$

and

$$\begin{aligned} (g(y(u)))'(v-u) &= \lim_{t \to 0^+} \frac{1}{t}\Big(g(y(u + t(v-u))) - g(y(u))\Big) \\ &= \lim_{t \to 0^+} \frac{1}{2t} E\Big(\int_D \big[|y(u + t(v-u)) - y(u)|^2 + 2(y(u+t(v-u)) - y(u), y - z_d)\big]\Big) \\ &= E\Big(\int_D y'(u)(v-u) \cdot (y - z_d) dx\Big) = [y'(u)(v-u), y - z_d]. \end{aligned} \tag{24}$$

Next, let us differentiate the state equation (36) at u in the direction v.
By (36), we have

$$\frac{1}{t}\big\{A[y(u+t(v-u)-y(u), w] - G[y(u)(u+t(v-u)) - y(u), w]\big\} = [v-u, w], \quad \forall\, w \in Y. \tag{25}$$

Taking limits in (25) as $t \to 0$, we obtain

$$A[y'(u)(v-u), w] - G[y'(u)(v-u), w] = [v-u, w], \quad \forall\, v \in K, \quad w \in Y. \tag{26}$$

Define the co-state $p \in Y$ satisfying

$$A[q, p] - G[q, p] = [y - z_d, q], \ \forall\, q \in Y, \tag{27}$$

here the existence of p is shown in Theorem 2.1.
Letting $w = p$ in (26), we have

$$[v-u, p] = A[y'(u)(v-u), p] - G[y'(u)(v-u), p] = [y'(u)(v-u), y - z_d] = (g(y(u)))'(v-u). \tag{28}$$

By (22)-(23) and (28), the optimality condition reads

$$\mathcal{J}'(u)(v-u) = [\alpha u + p, v - u] \geqslant 0, \qquad \forall\, v \in K, \tag{29}$$

where p is defined in (27). This completes the proof of Theorem2.3. □

It is known that the optimality conditions in (21) are necessary and sufficient.

3. Finite dimensional truncations of random fields

3.1. Finite expansion of random fields. In this paper, we will assume that the source of randomness can be expressed by a finite number of random variables that are mutually independent, and that we have finite expansions of input data. For those reasons, following the theory of Babuska [3], Wiener [58], as well as Xiu and Karniadakis [59], we can employ the following finite-dimensional noise assumption

Assumption 3.1. (finite dimensional noise) Any general second-order random process $X(\omega)$, $\omega \in \Omega$ can be represented in terms of a prescribed finite number of random variables $\xi = \xi(\omega) = (\xi_1(\omega), \cdots, \xi_N(\omega))$ with independent components $\xi_i(\omega)$, $i = 1, \cdots, N \in \mathbb{N}$. Let $\Gamma_i = \xi_i(\Omega) \in \mathbb{R}$ be a bounded interval for $i = 1, \cdots, N$ and $\rho_i : \Gamma_i \to [0, 1]$ be the probability density functions of the random variables $\xi_i(\omega)$, $\omega \in \Omega$. Then we can use the joint probability density function $\rho(\xi) = \prod_{i=1}^N \rho_i(\xi_i)$ for random vector ξ with the support $\Gamma = \prod_{i=1}^N \Gamma_i \subset \mathbb{R}^N$. On Γ, we have the probability measure $\rho(\xi)d\xi$.

As an example, we can use a finite-term expansion of the stochastic coefficient a and g based on N random variables (cf. [48]) :

$$a(x, \xi) = \sum_{i=1}^{S} \alpha_i(x) L_i(\xi), \;\; x \in D, \; \xi \in \Gamma, \tag{30}$$

$$g(s, x, \xi) = \sum_{i=1}^{S} \gamma_i(s, x) L_i(\xi), \;\; s, x \in D, \; \xi \in \Gamma, \tag{31}$$

where $\alpha_i(x) : D \to \mathbb{R}$, $\gamma_i(s, x) : D \times D \to \mathbb{R}$ and $L_i : \Gamma \to \mathbb{R}$. If a and g are represented by truncated Karhunen-Loève (KL) expansions [3], then $S = N + 1$ with $L_i = \xi_{i-1}$ and $\xi_0 = 1$; if a generalized polynomial chaos expansion [59] is used, L_i is an $N-$variate polynomial of order p and $S = (N + p)!/(N!p!)$.

As commented in [59], the above finite-term expansion allows us to conduct numerical formulations in the finite dimensional (N-dimensional) random space Γ. Let us denote $L^2_\rho(\Gamma)$ as the probabilistic Hilbert space [40], in which the random processes based upon the random variables ξ reside. The inner product of this Hilbert space is given by

$$\langle X, Y \rangle_{L^2_\rho(\Gamma)} = \int_\Gamma X(\xi) Y(\xi) \rho(\xi) d\xi, \;\; \forall \; X, \; Y \in L^2_\rho(\Gamma),$$

where we have exploited independence of the random variables to allow us to write the measure as product of measures in each stochastic direction. We similarly define the expectation of a random process $X \in L^2_\rho(\Gamma)$ as

$$\mathbb{E}[X(\xi)] = \int_\Gamma X(\xi) \rho(\xi) d\xi.$$

Additionally, we define the mapping $f : (x, \xi) \in D \times \Gamma \to \mathbb{R}$ be a set of random processes, which are indexed by the spatial position $x \in D$. Such a set of processes is referred to as a random field [25] and can also be interpreted as a function-valued random variable, because for every $\xi \in \Gamma$ the realization $f(\cdot, \xi) : D \to \mathbb{R}$ is a real valued function on D.

For a vector-space W on D, let the class $L^2_\rho(\Gamma; W)$ denote the space of random fields whose realizations lie in W for a.e (almost every) $\xi \in \Gamma$. If W is a Banach space, a norm on $L^2_\rho(\Gamma;W)$ is induced by $||f(x,\xi)||^2_{L^2_\rho(\Gamma;W))} = \mathbb{E}[\|f(x,\xi)\|^2_W]$; for example, on $L^2_\rho(\Gamma;L^2(D))$ we have

$$||f(x,\xi)||^2_{L^2_\rho(\Gamma;L^2(D))} = \mathbb{E}[\|f(x,\xi)\|^2_{L^2(D)}] = \int_\Gamma\int_D (f(x,\xi))^2\rho(\xi)dxd\xi,$$

which denotes the expected value of the $L^2(D)$-norm of the function $f(x,\xi)$.

We now give a Banach space that will be used as the solution space for the stochastic optimality system of equations, cf. [10]. Here, a Banach space $C_\rho(\Gamma;H)$ comprises all continuous functions $f:\Gamma\to H$ with a norm $\|f\|_{C_\rho(\Gamma;H)} \equiv \sup_{\xi\in\Gamma}\|f(\cdot,\xi)\|_H$, where H is a Hilbert space. Similarly, $C^p_\rho(\Gamma;H)$ is a Banach space with a norm

$$\|f\|_{C^p_\rho(\Gamma;H)} = \|f\|_{C_\rho(\Gamma;H)} + \sum_{j=1}^{N}\sum_{k=1}^{p_j}\|\partial^k_{\xi_j} f\|_{C_\rho(\Gamma;H)},$$

where $p=(p_1,p_2,...,p_N)$.

3.2. Finite dimensional representation of control problem. It follows from Doob-Dynkin theory (cf. [44]) that the solution y corresponding to the stochastic partial differential equation (2), can be described by just a finite number of random variables, i.e., $y(x,\omega)=y(x,\xi)=y(x,\xi_1(\omega),\cdots,\xi_N(\omega))$. The number N has to be large enough so that the approximation error is sufficient small.

Thus the stochastic optimal control problem (1)-(2) can be recast as a deterministic PDE-constrained optimization problem as follows:

$$\min_{u\in K}\mathcal{J}(u) = \min_{u\in K}\mathbb{E}\big(\frac{1}{2}\int_D|y-y_d|^2dx + \frac{\alpha}{2}\int_D|u|^2dx\big) \tag{32}$$

subject to

$$\begin{cases} -\nabla\cdot[a(x,\xi)\nabla y(x,\xi)] - \int_D g(s,x,\xi)y(s,\xi)ds = u(x), & x\in D,\ \xi\in\Gamma,\\ y(x,\xi)=0, & x\in\partial D,\ \xi\in\Gamma.\end{cases} \tag{33}$$

Here, we will replace the aforementioned assumption (3) by

$$\begin{aligned} c\|v\|^2_{H^1(D)} &\leqslant \big(a(x,\xi)\nabla v,\nabla v\big)_{L^2(D)} - \big(\int_D g(s,x,\xi)v(s,\xi)ds, v(x)\big)_{L^2(D)}\\ &\leqslant C\|v\|^2_{H^1(D)},\ \forall\ v\in H^1(D),\ \text{a.e. }\Gamma. \end{aligned} \tag{34}$$

Using these finite dimensional representation the weak formulation of deterministic optimal control problem (32)-(33) can be rewritten as:

$$\min_{u\in K}\mathcal{J}(u) = \min_{u\in K}\mathbb{E}\big(\frac{1}{2}\int_D|y-y_d|^2dx + \frac{\alpha}{2}\int_D|u|^2dx\big) \tag{35}$$

subject to

$$A[y,v] - G[y,v] = [u,v],\quad \forall\ v\in Y_\rho, \tag{36}$$

where we have taken the deterministic state space $Y_\rho = L^2_\rho(\Gamma;H^1_0(D))$, and

$$A[y,v] = \int_\Gamma\int_D a(x,\xi)\nabla y(x,\xi)\cdot\nabla v(x,\xi)\rho(\xi)dxd\xi = \int_\Gamma\int_D a\nabla y\cdot\nabla v\rho dxd\xi,\ \forall y,v\in Y_\rho, \tag{37}$$

(38)
$$G[y,v] = \int_\Gamma \int_D \big(\int_D g(s,x,\xi) y(s,\xi) v(x,\xi) \rho(\xi) ds \big) dx d\xi$$
$$= \int_\Gamma \int_D \big(\int_D g(s,x) y(s) v(x) \rho(\xi) ds \big) dx d\xi, \qquad \forall y,, v \in Y_\rho,$$

and
(39)
$$[u,v] = \int_\Gamma \int_D u(x,\xi) v(x,\xi) \rho(\xi) dx d\xi = \int_\Gamma \int_D uv\rho dx d\xi, \quad \forall\ u,\ v \in L^2(\Gamma; L^2(D)).$$

Under the assumption (34), the existence of solutions to (35)-(36) can be proved similarly.

It follows from [13, 33] that the optimal control problem (35)-(36) has a unique solution $(y,u) \in Y_\rho \times K$. Furthermore, a pair (y,u) is the solution of (35)-(36) iff there is a co-state variable $p \in Y_\rho$, such that the triplet (y,p,u) satisfies the following optimality system:

$$\begin{cases} A[y,v] - G[y,v] = [u,v], \quad \forall\ v \in Y_\rho, \\ A[q,p] - G[q,p] = [y - y_d, q], \quad \forall\ q \in Y_\rho, \\ [p + \alpha u, w - u] \geqslant 0, \quad \forall\ w \in K \subset U. \end{cases} \tag{40}$$

It is known that the inequality in (40) is just the necessary and sufficient optimality condition.

The explicit solution of the variational inequality in (40) depends heavily on the choice of the joint probability density ρ. If the joint probability density ρ is uniform on Γ, we have the following explicit solution for some cases (see, e.g. [33, 36]). For example

I: Let K be given by

$$K = \{u \in L^2(D) :\ u(x) \geqslant d,\ \forall\ x \in D\}. \tag{41}$$

Then, the solution is

$$u(x) = max\big\{d, -\frac{1}{\alpha}\mathbb{E}[p(x,\xi)]\big\}, \quad \text{a.e.}\ \ (x,\xi) \in D \times \Gamma. \tag{42}$$

II: Let K be given by

$$K = \{u \in L^2(D) :\ \int_D u(x) \geqslant d,\ \forall\ x \in D\}. \tag{43}$$

Then, the solution is

$$u(x) = -\frac{1}{\alpha}\mathbb{E}[p(x,\xi)] + max\big\{d, \frac{1}{\alpha}\overline{\mathbb{E}p}\big\}, \quad \text{a.e.}\ \ (x,\xi) \in D \times \Gamma, \tag{44}$$

where $\overline{\mathbb{E}p} = \dfrac{\int_D \mathbb{E}[p(x,\xi)] dx}{\int_D dx}$.

III: Let K be given by

$$K = \{u \in L^2(D) :\ c \leqslant u(x) \leqslant d,\ \forall\ x \in D\}, \tag{45}$$

where constants c, $d \in \mathbb{R}$ and $c < d$. Then, the solution is

$$u(x) = \begin{cases} c, & \text{if}\ \ \mathbb{E}[p(x,\xi)] + \alpha u(x) > 0, \\ -\frac{1}{\alpha}\mathbb{E}[p(x,\xi)], & \text{if}\ \ \mathbb{E}[p(x,\xi)] + \alpha u(x) = 0, \\ d, & \text{if}\ \ \mathbb{E}[p(x,\xi)] + \alpha u(x) < 0, \end{cases} \quad \text{a.e.}\ (x,\xi) \in D \times \Gamma. \tag{46}$$

Also, we can rewrite the solution as

$$u(x) = Proj_{[c,d]}\{-\frac{1}{\alpha}\mathbb{E}[p(x,\xi)]\}, \text{ a.e. } (x,\xi) \in D \times \Gamma, \tag{47}$$

where $Proj_{[c,d]}$ denotes the projection mapping from $\mathbb{R}$ onto $[c,d]$.

4. Stochastic Galerkin method: stochastic finite element on $D \times \Gamma$

In this work we adopt the stochastic finite element spaces defined on $D \times \Gamma$ by [4, 23].

We firstly consider finite element spaces defined on domain $D \subset \mathbb{R}^d$. Let $\{\mathcal{T}_h\}_{h>0}$ be a family of regular triangulation of D such that $\bar{D} = \cup_{\tau \in \mathcal{T}_h} \bar{\tau}$. Let $h_s = \max_{\tau \in \mathcal{T}_h} h_\tau$, where h_τ denotes the diameter of the element τ. Consider two finite element spaces $V_{h_s} \subset H_0^1(D)$ and $W_{h_s} \subset L^2(D)$, consisting of piecewise linear (or constant) continuous functions on $\{\mathcal{T}_h\}$, respectively. Here V_{h_s} and W_{h_s} satisfy the following approximation properties [7]:

(i) for all $\phi \in H^2(D) \cap H_0^1(D)$, there exists

$$\inf_{\phi_{h_s} \in V_{h_s}} \|\phi - \phi_{h_s}\|_{H_0^1(D)} \leqslant C h_s \|\phi\|_{H^2(D)}, \tag{48}$$

where $C > 0$ is a constant independent of ϕ and h_s.

(ii) for all $\phi \in H_0^1(D)$, there exists

$$\inf_{\phi_{h_s} \in W_{h_s}} \|\phi - \phi_{h_s}\|_{L^2(D)} \leqslant C h_s \|\phi\|_{H_0^1(D)}, \tag{49}$$

where $C > 0$ is a constant independent of ϕ and h_s.

Then, we establish a finite dimensional space defined on $\Gamma \subset \mathbb{R}^N$ ([3]). Let Γ be partitioned into a finite number of disjoint boxes $B_i^N \subset \mathbb{R}^N$, that is, for a finite index set I, we have

$$\Gamma = \bigcup_{i \in I} B_i^N = \bigcup_{i \in I} \prod_{j=1}^{N} (a_i^j, \, b_i^j),$$

where $B_k^N \cap B_l^N = \varnothing$ for $k \neq l \in I$ and $(a_i^j, b_i^j) \subset \Gamma_j$. A maximum grid size parameter $0 < h_r < 1$ is denoted by

$$h_r = \max\{|b_i^j - a_i^j|/2 : 1 \leqslant j \leqslant N \,\text{and}\, i \in I\}.$$

Let $S_{h_r} \subset L_\rho^2(\Gamma)$ be the finite element space of piecewise polynomials with degree at most p_j on each direction ξ_j. Thus if $\psi_{h_r} \in S_{h_r}$, then $\psi_{h_r}|_{B_i^N} \in \text{span}\{\Pi_{j=1}^N \xi_j^{n_j} : n_j \in \mathbb{N} \,\text{and}\, n_j \leqslant p_j\}$. Letting the multi-index $p = (p_1, \cdots, p_N)$, we have (see [6]) the following property: for all $\psi \in C^{p+1}(\Gamma)$,

$$\inf_{\psi_{h_r} \in S_{h_r}} \|\psi - \psi_{h_r}\|_{L_\rho^2(\Gamma)} \leqslant h_r^\gamma \sum_{j=1}^{N} \frac{\|\partial_{\xi_j}^{p_j+1} \psi\|_{L_\rho^2(\Gamma)}}{(p_j+1)!}, \tag{50}$$

where $\gamma = \min_{1 \leqslant j \leqslant N} \{p_j + 1\}$.

Combining spaces V_{h_s}, W_{h_s} and S_{h_r} together, one can obtain a tensor product finite element space on $D \times \Gamma$.

We will use $Y_h = V_{h_s} \times S_{h_r}$ to approximate the sate variable y and co-state variable p, $U_h = W_{h_s}$ to approximate the control. Let $K_h = U_h \cap K$ be the approximation of the admissible set K.

Now define the $H_0^1(D)$-projection operator R_{h_s} as follows: $H_0^1(D) \to V_{h_s}$ by

$$(R_{h_s}\phi, \phi_{h_s})_{H_0^1(D)} = (\phi, \phi_{h_s})_{H_0^1(D)}, \quad \forall\, \phi_{h_s} \in V_{h_s}, \ \forall\, \phi \in H_0^1(D), \tag{51}$$

the $L^2(D)$-projection operator Π_{h_s}: $L^2(D) \to W_{h_s}$ by

$$(\Pi_{h_s}\phi, \phi_{h_s})_{L^2(D)} = (\phi, \phi_{h_s})_{L^2(D)}, \quad \forall\ \phi_{h_s} \in W_{h_s},\ \forall\ \phi \in L^2(D). \tag{52}$$

In later theory derivation it is important to have $\Pi_{h_s}K \subseteq K_h$. However this will generally depend on the structure of K and the finite element spaces chosen. Here we only show that if we use the piecewise constant space for W_{h_s} , then we have $\Pi_{h_s}K \subseteq K$.

For example, we show the proof of Case III for K. The proofs of other cases are similar. Notice that

$$K = \{u \in L^2(D) :\ c \leqslant u(x) \leqslant d,\ \forall\ x \in D\}, \tag{53}$$

where constants c, $d \in \mathbb{R}$ and $c < d$. If we choose $\phi_{h_s} = 1$ in (52), we can see that $\int_\tau \Pi_{h_s}u = \int_\tau u$. Noticing that W_{h_s} is piecewise constant, we know

$$c\tau \leqslant \tau\Pi_{h_s}u = \int_\tau \Pi_{h_s}u = \int_\tau u \leqslant d\tau. \tag{54}$$

Thus $c \leqslant \Pi_{h_s}u \leqslant d$. This means $\Pi_{h_s}K \subseteq K_h$.

Similarly, define the $L^2_\rho(\Gamma)$-projection operator Π_{h_r} as follows: $L\rho^2(\Gamma) \to S_{h_r}$ by

$$(\Pi_{h_r}\psi, \psi_{h_r})_{L\rho^2(\Gamma)} = (\psi, \psi_{h_r})_{L\rho^2(\Gamma)}, \quad \forall\ \psi_{h_r} \in S_{h_r},\ \forall\ \psi \in L\rho^2(\Gamma). \tag{55}$$

From (48), for all $\phi \in H^2(D) \cap H^1_0(D)$

$$\|\phi - R_{h_s}\phi\|_{H^1_0(D)} \leqslant Ch_s\|\phi\|_{H^2(D)}, \tag{56}$$

and from (49) that for all $\phi \in H^1(D)$

$$\|\phi - \Pi_{h_s}\phi\|_{L^2(D)} \leqslant Ch_s\|\phi\|_{H^1(D)}. \tag{57}$$

It follows from (50) that for all $\psi \in C^{p+1}_\rho(\Gamma)$

$$\|\psi - \Pi_{h_r}\psi\|_{L^2_\rho(\Gamma)} \leqslant h_r^\gamma \sum_{j=1}^{N} \frac{\|\partial^{p_j+1}_{\xi_j}\psi\|_{L^2_\rho(\Gamma)}}{(p_j+1)!}. \tag{58}$$

It follows from the inequalities (56) and (58) that we have the following approximation property (cf.[3], Proposition 3.1): for all $y \in C^{p+1}_\rho\big(\Gamma; H^2(D) \cap H^1_0(D)\big)$

(59)

$$\inf_{y_h \in Y_h} \|y - y_h\|_{L^2_\rho(\Gamma;H^1_0(D))} \leqslant C\big\{h_s\|y\|_{L^2_\rho(\Gamma;H^2(D))} + h_r^\gamma \sum_{j=1}^{N} \frac{\|\partial^{p_j+1}_{\xi_j}y\|_{L^2_\rho(\Gamma;H^1_0(D))}}{(p_j+1)!}\big\},$$

where positive constant C is independent of h_s, h_r, N and p.

Next we introduce a projection operator P_h that maps onto the tensor product space $U_h = W_{h_s} \times S_{h_r}$ as follows:

$$P_h\varphi = \Pi_{h_s}\Pi_{h_r}\varphi = \Pi_{h_r}\Pi_{h_s}\varphi, \quad \forall\ \varphi \in\ L^2_\rho(\Gamma; L^2(D)). \tag{60}$$

Also we need the following decomposition

$$\varphi - P_h\varphi = (\varphi - \Pi_{h_s}\varphi) + \Pi_{h_s}(I - \Pi_{h_r})\varphi, \quad \forall\ \varphi \in\ L^2_\rho(\Gamma; L^2(D)). \tag{61}$$

The following assumption and lemmas on the regularity are needed in deriving the error estimates later.

Assumption 4.1. Let y, p, u satisfy the following regularity condition

$$y,\ p \in C^{p+1}\big(\Gamma; H^2(D) \cap H^1_0(D)\big),\ u \in H^1(D). \tag{62}$$

Lemma 4.1. [18] *Let $u \in L^2(D)$. Then for any $\xi \in \Gamma$, $y(\cdot,\xi) \in H^2(D)$ and there exists $C > 0$ such that*

$$\|y(\cdot,\xi)\|_{H^2(D)} \leqslant C\|u\|_{L^2(D)}. \tag{63}$$

Similar to Lemma 3.7 in [19], the following lemma will be also used. It can be proved by an inductive argument after taking derivatives with respect to ξ_j of (33) and using the Greens formulas.

Lemma 4.2. *Let $u \in L^2(D)$ and $\varphi_j \in L^\infty(D)$. Then for all $j = 1, 2, \cdots, N$ and for any $\xi \in \Gamma$, there exists $C > 0$ such that*

$$\frac{\|\partial_{\xi_j}^{p_j+1} y(\cdot,\xi)\|_{H_0^1(D)}}{(p_j+1)!} \leqslant C\|\varphi_j\|_{L^\infty(D)}^{p_j+1}\|u\|_{L^2(D)}. \tag{64}$$

5. Galerkin approximation scheme and a priori error estimates

In this section we apply the stochastic Galerkin finite element to approximate the optimal control problem (35)-(36), as formulated below:

$$\min_{u_h \in K_h} \mathcal{J}_h(u_h) = \min_{u_h \in K_h} \mathbb{E}\Big(\frac{1}{2}\int_D |y_h - y_d|^2 dx + \frac{\alpha}{2}\int_D |u_h|^2 dx\Big) \tag{65}$$

subject to

$$A[y_h, v_h] - G[y_h, v_h] = [u_h, v_h], \quad \forall\, v_h \in Y_h. \tag{66}$$

Similarly, we have (e.g., see [13, 33]) that the control problem (65)-(66) has a unique pair solution $(y_h, u_h) \in Y_h \times K_h$, if and only if there is a co-state variable $p_h \in Y_h$, such that $\{y_h,\, p_h,\, u_h\} \in Y_h \times Y_h \times K_h$ satisfies the following system

$$\begin{cases} A[y_h, v_h] - G[y_h, v_h] = [u_h, v_h], \quad \forall\, v_h \in Y_h, \\ A[q_h, p_h] - G[q_h, p_h] = [y_h - y_d, q_h], \quad \forall\, q_h \in Y_h, \\ [p_h + \alpha u_h, w_h - u_h] \geqslant 0, \quad \forall\, w_h \in K_h \subset U_h. \end{cases} \tag{67}$$

As in the continuous case of (20), the discrete directional derivative of functional $\mathcal{J}$ reads:

$$\mathcal{J}_h'(u_h)(w_h) = [p_h, w_h] + \alpha[u_h, w_h], \quad \forall\, w_h \in K_h \subset U_h, \tag{68}$$

$$\mathcal{J}_h'(u_h)(w_h - u_h) \geqslant 0, \quad \forall\, w_h \in K_h \subset U_h. \tag{69}$$

In the follows we will derive a priori error estimates for the stochastic Galerkin scheme. To this end we need to firstly introduce the following intermediate variational inequality for the semi-discrete problem where u is is not discrete. We have

$$\mathcal{J}_h'(u)(w-u) = [p_h(u), w-u] + \alpha[u, w-u] \geqslant 0, \quad \forall\, w \in K \subset U, \tag{70}$$

where $p_h(u) \in Y_h$ is the solution of the system:

$$\begin{cases} A[y_h(u), v_h] - G[y_h(u), v_h] = [u, v_h], \quad \forall\, v_h \in Y_h, \\ A[q_h, p_h(u)] - G[q_h, p_h(u)] = [y_h(u) - y_d, q_h], \quad \forall\, q_h \in Y_h. \end{cases} \tag{71}$$

We are now in the position of deriving a priori error estimates for our scheme. To this end we need a series of lemmas, which will be proved below. Firstly we have:

Lemma 5.1. *Under the definition of* (70)*, the following inequality holds:*

$$\mathcal{J}_h'(w)(w-u) - \mathcal{J}_h'(u)(w-u) \geqslant \alpha\|w-u\|^2_{L^2_\rho(\Gamma;L^2(D))}. \tag{72}$$

Proof. By (70), we have

$$\mathcal{J}_h'(w)(w-u)-\mathcal{J}_h'(u)(w-u)=[p_h(w)-p_h(u),w-u]+\alpha[w-u,w-u]. \tag{73}$$

Furthermore from (71) we have

$$\begin{aligned}&[p_h(w)-p_h(u),w-u]\\ =&A[y_h(w)-y_h(u),p_h(w)-p_h(u)]-G[y_h(w)-y_h(u),p_h(w)-p_h(u)]\\ =&[y_h(w)-y_h(u),y_h(w)-y_h(u)].\end{aligned} \tag{74}$$

Hence, from (73) and (74) we can infer (72). □

Lemma 5.2. *Let* (y,p,u) *be the solution of the optimal control problem* (21) *and* (y_h,p_h,u_h) *be the solution of the discretized problem* (67). *Let the Assumption 4.1 be fulfilled and* $\Pi_{h_s}K\subseteq K_h$. *Then the following estimate holds:*

$$\begin{aligned}\|u-u_h\|_{L^2_\rho(\Gamma;L^2(D))}\leqslant&\|p-p_h(u)\|_{L^2_\rho(\Gamma;L^2(D))}+Ch_s\big\{\|u\|_{L^2_\rho(\Gamma;H^1(D))}+\|p\|_{L^2_\rho(\Gamma;H^1(D))}\big\}\\ &+Ch_r^\gamma\sum_{j=1}^N\frac{\|\partial_{\xi_j}^{p_j+1}p\|_{L^2_\rho(\Gamma;L^2(D))}}{(p_j+1)!},\end{aligned} \tag{75}$$

where $\gamma=\min\limits_{1\leqslant j\leqslant N}\{p_j+1\}$.

Proof. It follows from (69), (70) and Lemma 5.1 that

$$\begin{aligned}\alpha\|u-u_h\|^2_{L^2_\rho(\Gamma;L^2(D))}&\leqslant\mathcal{J}_h'(u)(u-u_h)-\mathcal{J}_h'(u_h)(u-u_h)\\ &=[\alpha u+p_h(u),u-u_h]-[\alpha u_h+p_h,u-u_h]\\ &=[\alpha u+p,u-u_h]-[p-p_h(u),u-u_h]\\ &\quad+[\alpha u_h+p_h,u_h-\Pi_{h_s}u]+[\alpha u_h+p_h,\Pi_{h_s}u-u]\\ &\leqslant[p_h(u)-p,u-u_h]+[\alpha u_h+p_h,\Pi_{h_s}u-u].\end{aligned} \tag{76}$$

Note that

$$\begin{aligned}&[\alpha u_h+p_h,\Pi_{h_s}u-u]\\ =&[\alpha u+p,\Pi_{h_s}u-u)]+\big[\alpha(u_h-u),\Pi_{h_s}u-u\big]\\ &+[p_h-p_h(u),\Pi_{h_s}u-u]+[p_h(u)-p,\Pi_{h_s}u-u]\\ &\leqslant[\alpha u+p,\Pi_{h_s}u-u]+C(\delta)\|u-\Pi_{h_s}u\|^2_{L^2_\rho(\Gamma;L^2(D))}+C\delta\|u-u_h\|^2_{L^2_\rho(\Gamma;L^2(D))}\\ &+C\delta\|p_h-p_h(u)\|^2_{L^2_\rho(\Gamma;L^2(D))}+C\delta\|p_h(u)-p\|^2_{L^2_\rho(\Gamma;L^2(D))},\end{aligned} \tag{77}$$

where δ is an arbitrary small positive number. Furthermore

$$\|u-\Pi_{h_s}u\|_{L^2_\rho(\Gamma;L^2(D))}\leqslant Ch_s\|u\|_{L^2_\rho(\Gamma;H^1(D))}, \tag{78}$$

and

(79)

$$\begin{aligned}&[\alpha u+p,\Pi_{h_s}u-u]=[\alpha u+p-P_h(\alpha u+p),\Pi_{h_s}u-u]\\ &\leqslant C\big\{h_s^2\|u\|^2_{L^2_\rho(\Gamma;H^1(D))}+C\big\{h_s^2\|p\|^2_{L^2_\rho(\Gamma;H^1(D))}+h_r^{2\gamma}\sum_{j=1}^N\Big(\frac{\|\partial_{\xi_j}^{p_j+1}p\|_{L^2_\rho(\Gamma;L^2(D))}}{(p_j+1)!}\Big)^2\big\}.\end{aligned}$$

Then from (76)-(79) we have

(80)
$$\begin{aligned}&\alpha\|u-u_h\|^2_{L^2_\rho(\Gamma;L^2(D))}\\ \leqslant\ & C\|p-p_h(u)\|^2_{L^2_\rho(\Gamma;L^2(D))}+C\delta\|u-u_h\|^2_{L^2_\rho(\Gamma;L^2(D))}+C\delta\|p_h-p_h(u)\|^2_{L^2_\rho(\Gamma;L^2(D))}\\ &+C\big\{h_s^2\|u\|^2_{L^2_\rho(\Gamma;H^1(D))}+h_s^2\|p\|^2_{L^2_\rho(\Gamma;H^1(D))}+h_r^{2\gamma}\sum_{j=1}^{N}\big(\frac{\|\partial_{\xi_j}^{p_j+1}p\|_{L^2_\rho(\Gamma;L^2(D))}}{(p_j+1)!}\big)^2\big\}.\end{aligned}$$

It follows from (67) and (71) that

$$\|p_h-p_h(u)\|_{L^2_\rho(\Gamma;H^1(D))}\leqslant C\|y_h-y_h(u)\|_{L^2_\rho(\Gamma;H^1(D))}\leqslant C\|u-u_h\|_{L^2_\rho(\Gamma;L^2(D))}. \quad (81)$$

Then (75) follows from (80) and (81). This completes the proof of Lemma 5.2. □

Now, we give an estimate for the error in the adjoint states

Lemma 5.3. *Let* (y,p,u) *be the solution of the optimal control problem* (21) *and* $(y_h(u),p_h(u),u)$ *be the solution of the auxiliary problem* (71). *Then the following estimates hold:*

$$||y-y_h(u)||_{L^2_\rho(\Gamma;H^1(D))}\leqslant Ch_s\|y\|_{L^2_\rho(\Gamma;H^2(D))}, \quad (82)$$

and

$$||p-p_h(u)||_{L^2_\rho(\Gamma;H^1(D))}\leqslant Ch_s\big\{\|y\|_{L^2_\rho(\Gamma;H^2(D))}+\|p\|_{L^2_\rho(\Gamma;H^2(D))}\big\}. \quad (83)$$

Proof. Let us firstly note the following Galerkin orthogonality equation from (21) and (71):

$$A[y-y_h(u),v_h]-G[y-y_h(u),v_h]=0,\quad \forall\ v_h\in Y_h. \quad (84)$$

Now split the error $y-y_h(u)=(y-R_{h_s}y)+(R_{h_s}y-y_h(u))$, where the $H^1(D)$-projection R_{h_s} is given by (51). It follows from the elliptic condition (34) and (84) that

$$\begin{aligned}c||y-y_h(u)||^2_{L^2_\rho(\Gamma;H^1(D))}&\leqslant A[y-y_h(u),y-y_h(u)]-G[y-y_h(u),y-y_h(u)]\\ &=A[y-y_h(u),y-R_{h_s}y]-G[y-y_h(u),y-R_{h_s}y].\end{aligned} \quad (85)$$

It follows from the approximation property (56) that

$$||y-y_h(u)||_{L^2_\rho(\Gamma;H^1(D))}\leqslant Ch_s\|y\|_{L^2_\rho(\Gamma;H^2(D))}. \quad (86)$$

Similarly there holds the another Galerkin orthogonality equation from the adjoint equation:

$$A[q_h,p-p_h(u)]-G[q_h,p-p_h(u)]=[y-y_h(u),q_h],\ \forall\ q_h\in Y_h. \quad (87)$$

Again split the error $p-p_h(u)=(p-p_I)+(p_I-p_h(u))$, where $p_I\in V_{h_s}$ is the Lagrange interpolation of p. It follows from the elliptic condition (34) and (87) that

$$\begin{aligned}&c||p-p_h(u)||^2_{L^2_\rho(\Gamma;H^1(D))}\\ \leqslant\ &A[p-p_h(u),p-p_h(u)]-G[p-p_h(u),p-p_h(u)]\\ =\ &A[p-p_I,p-p_h(u)]+A(p_I-p_h(u),p-p_h(u)]-G[p-p_I,p-p_h(u)]\\ &-G[p_I-p_h(u),p-p_h(u)]\\ \leqslant\ &C\big\{\|p-p_I\|_{L^2_\rho(\Gamma;H^1(D))}+\|y-y_h(u)\|_{L^2_\rho(\Gamma;L^2(D))}\big\}||p-p_h(u)||^2_{L^2_\rho(\Gamma;H^1(D))}.\end{aligned} \quad (88)$$

Then it follows from the standard approximation property of Lagrange interpolation and (86) that

$$\begin{aligned} ||p - p_h(u)||_{L^2_\rho(\Gamma;H^1(D))} &\leqslant ||p - p_I||_{L^2_\rho(\Gamma;H^1(D))} + ||y - y_h(u)||_{L^2_\rho(\Gamma;L^2(D))} \\ &\leqslant Ch_s\big\{\|y\|_{L^2_\rho(\Gamma;H^2(D))} + \|p\|_{L^2_\rho(\Gamma;H^2(D))}\big\}. \end{aligned} \tag{89}$$

Then, we complete the proof of Lemma 5.3. □

It follows from Lemmas 5.2 and 5.3 that the following error estimates with respect to $y - y_h,\ p - p_h$ and $u - u_h$ hold.

Theorem 5.4. *Let* (y, p, u) *be the solution of the optimal control problem* (21) *and* (y_h, p_h, u_h) *be the solution of the discretized problem* (67)*, respectively. Assume that the conditions of Lemmas 5.1 to 5.3 are valid. Then the following error estimate holds:*

$$\begin{aligned} &\|y - y_h\|_{L^2_\rho(\Gamma;H^1(D))} + \|p - p_h\|_{L^2_\rho(\Gamma;H^1(D))} + \|u - u_h\|_{L^2_\rho(\Gamma;L^2(D))} \\ \leqslant &Ch_s\big\{\|y\|_{L^2_\rho(\Gamma;H^2(D))} + \|p\|_{L^2_\rho(\Gamma;H^2(D))} + \|u\|_{L^2_\rho(\Gamma;H^1(D))}\big\} \\ &+ Ch_r^\gamma \sum_{j=1}^{N} \frac{\|\partial_{\xi_j}^{p_j+1} p\|_{L^2_\rho(\Gamma;L^2(D))}}{(p_j+1)!}, \end{aligned} \tag{90}$$

where $\gamma = \min\limits_{1\leqslant j\leqslant N}\{p_j + 1\}$.

Proof. From (75) and (83) we have

$$\begin{aligned} &\|u - u_h\|_{L^2_\rho(\Gamma;L^2(D))} \\ \leqslant\ &Ch_s\big\{\|y\|_{L^2_\rho(\Gamma;H^2(D))} + \|p\|_{L^2_\rho(\Gamma;H^2(D))} + \|u\|_{L^2_\rho(\Gamma;H^1(D))}\big\} \\ &+Ch_r^\gamma \sum_{j=1}^{N} \frac{\|\partial_{\xi_j}^{p_j+1} p\|_{L^2_\rho(\Gamma;L^2(D))}}{(p_j+1)!}, \end{aligned} \tag{91}$$

where $\gamma = \min\limits_{1\leqslant j\leqslant N}\{p_j + 1\}$.

Furthermore, by using (81), (83) and (91) we have

$$\begin{aligned} &\|y - y_h\|_{L^2_\rho(\Gamma;H^1(D))} + \|p - p_h\|_{L^2_\rho(\Gamma;H^1(D))} \\ \leqslant\ &\|y - y_h(u)\|_{L^2_\rho(\Gamma;H^1(D))} + \|y_h(u) - y_h\|_{L^2_\rho(\Gamma;H^1(D))} \\ &+\|p - p_h(u)\|_{L^2_\rho(\Gamma;H^1(D))} + \|p_h(u) - p_h\|_{L^2_\rho(\Gamma;H^1(D))} \\ \leqslant\ &C\|u - u_h\|_{L^2_\rho(\Gamma;L^2(D))} + Ch_s\big\{\|y\|_{L^2_\rho(\Gamma;H^2(D))} + \|p\|_{L^2_\rho(\Gamma;H^2(D))}\big\} \\ \leqslant\ &Ch_s\big\{\|y\|_{L^2_\rho(\Gamma;H^2(D))} + \|p\|_{L^2(\Gamma;H^2(D))} + \|u\|_{L^2_\rho(\Gamma;H^1(D))}\big\} \\ &+Ch_r^\gamma \sum_{j=1}^{N} \frac{\|\partial_{\xi_j}^{p_j+1} p\|_{L^2_\rho(\Gamma;L^2(D))}}{(p_j+1)!}. \end{aligned} \tag{92}$$

Then (90) is obtained from (91) and (92). This completes the proof. □

6. Numerical experiments

In this section we study numerical algorithms resulted from the stochastic finite element approximation set up in Section 5. We firstly examine the characteristic of the matrix system of the finite approximation. Here we only consider the case where space S_{h_r} has no partition of Γ, i.e. the approximation is achieved by increasing the polynomial degree only. We use the tensor finite element space $S_{h_r} = \bigotimes_{n=1}^{N} Z_n^{p_n}$, where the one dimensional global polynomial subspaces reads $Z_n^{p_n} = \{v: \Gamma_n \to \mathbb{R}: v \in \text{span}(1, y_n, \cdots, y_n^{p_n})\}$, $n = 1, \cdots, N$.

Let $\{\varphi_i(x)\}$ be the a basis of the space V_{h_s}, $\{\psi_j(\xi)\}$ be a basis of the space S_{h_r}. Assume the solutions of the discrete optimality system:

$$\begin{cases} A[y_h, v_h] - G[y_h, v_h] = [u_h, v_h], \quad \forall\, v_h \in Y_h, \\ A[q_h, p_h] - G[q_h, p_h] = [y_h - y_d, q_h], \quad \forall\, q_h \in Y_h, \\ [p_h + \alpha u_h, w_h - u_h] \geqslant 0, \quad \forall\, w_h \in K_h \subset U_h. \end{cases} \tag{93}$$

are presented by

$$\begin{cases} y_h = \sum_{i,j} y_{ij}\varphi_i(x)\psi_j(\xi), \\ p_h = \sum_{i,j} p_{ij}\varphi_i(x)\psi_j(\xi), \\ u_h = \sum_{i} u_i\varphi_i(x). \end{cases} \tag{94}$$

The above system will be solved by a well-known iterative method as described in [36], where u_h is given an initial value. Then the state and co-state equations are solved in turns and the control u_h is updated the explicit solution formula of the inequality of (93). Similarly to the continuous case the explicit solution of the inequality in (93) depends on the choice of the joint probability density ρ, and the finite element base function used to approximate to u_h. If ρ is uniform on Γ and the finite element base function used to approximate u_h is piecewise constant, we can obtain the explicit discrete solution (see, e.g.[36, 51]). For example we have

$$u_h(x) = max\{d, -\frac{1}{\alpha|\tau|}\int_\tau E(p_h(x,\xi))dx\}, \quad \forall\, x \in \tau \subset \mathcal{T}_h \text{ ,a.e. } \xi \in \Gamma, \tag{95}$$

for $K = \{u \in L^2(D): u(x) \geqslant d, \forall\, x \in D\}$.

Now we wish to examine the matrix structure of the two equations. As an example, we show the matrix structure for the state equation in (93). Taking the test function $v_h = \varphi_l(x)\psi_k(\xi)$, we have

$$\begin{aligned} &\sum_{i,j}\Big(\int_\Gamma \rho(\xi)\psi_k(\xi)\psi_j(\xi)(a\nabla\varphi_i, \nabla\varphi_l)_{L^2(D)}d\xi \\ &- \int_\Gamma \rho(\xi)\psi_k(\xi)\psi_j(\xi)(\int_D g(s,x,\xi)\varphi_i(s)ds, \varphi_l(x))_{L^2(D)}d\xi\Big)y_{ij} \\ &= \sum_i\Big(\int_\Gamma \rho(\xi)\psi_k(\xi)(\varphi_i, \varphi_l)_{L^2(D)}d\xi\Big)u_i, \quad \forall\, k,\ l, \end{aligned} \tag{96}$$

which can be reformulated as

$$\sum_{i,j}\Big(\int_\Gamma \rho(\xi)\psi_k(\xi)\psi_j(\xi)K_{i,l}(\xi)d\xi\Big)y_{ij} = \sum_i\Big(\int_\Gamma \rho(\xi)\psi_k(\xi)M_{i,l}d\xi\Big)u_i, \quad \forall\, k,\ l, \tag{97}$$

where $K_{i,l}(\xi) = (a(\cdot,\xi)\nabla\varphi_i, \nabla\varphi_l)_{L^2(D)} - (\int_D g(s,x,\xi)\varphi_i(s)ds, \varphi_l(x))_{L^2(D)}$ and $M_{i,l} = (\varphi_i, \varphi_l)_{L^2(D)}$.

Furthermore when the diffusion coefficients a, g are expanded by finite terms (30), i.e. $a(x,\xi) = \sum_{t=1}^{S}\alpha_t(x)L_t(\xi), g(s,x,\xi) = \sum_{t=1}^{S}\gamma_t(s,x)L_t(\xi)$ the corresponding expression for the stiffness matrix reads

$$K_{i,l}(\xi) \equiv \int_D \Big(\sum_{t=1}^{S} \alpha_t(x) L_t(\xi)\Big) \nabla\varphi_i(x) \cdot \nabla\varphi_l(x)\, dx$$

$$(98) \qquad - \int_D \Big(\sum_{t=1}^{S} \int_D \gamma_t(s,x,\xi)\varphi_i(s) ds L_t(\xi)\Big) \cdot \varphi_l(x)\, dx.$$

Since $\psi_k \in S_{h_r} = \bigotimes_{n=1}^{N} Z_n^{p_n}$, we only need to take it as the product $\psi_k(\xi) = \prod_{r=1}^{N} \psi_{kr}(\xi_r)$, where $\psi_{kr}:\ \Gamma_r \to \mathbb{R}$ is a basis function of the subspace

$$Z_r^{p_r} = \text{span}\{1, \xi_r, \cdots, \xi_r^{p_r}\} = \text{span}\{\psi_{kr} : kr = 1, \cdots, p_r + 1\}.$$

Inserting this ψ_k into (97), we have

$$\sum_{i,j} \Big(\int_\Gamma \prod_{r=1}^{N} \rho_r(\xi_r)\psi_{kr}(\xi_r)\psi_{jr}(\xi_r) K_{i,l}(\xi) d\xi\Big) y_{ij}$$

$$(99) \qquad = \sum_i \Big(\int_\Gamma \prod_{r=1}^{N} \rho_r(\xi_r)\psi_{kr}(\xi_r) M_{i,l} d\xi\Big) u_i, \quad \forall\ k,\ l.$$

Using (98), we obtain the coefficient of y_{ij} as follows:

$$(100) \qquad \sum_{t=1}^{S} K_{i,l}^t \int_\Gamma L_t(\xi) \prod_{r=1}^{N} \rho_r(\xi_r)\psi_{kr}(\xi_r)\psi_{jr}(\xi_r) d\xi,$$

where

$$(101) \qquad K_{i,l}^t = \int_D \alpha_t(x)\nabla\varphi_i(x) \cdot \nabla\varphi_l(x) dx - \int_D \int_D \gamma_t(s,x)\varphi_i(s)\varphi_l(x) ds dx.$$

Similarly one can handle the other equations. Next, we should be able to solve y_{ij}, p_{ij} and u_i which are the coefficients of the solutions of the discrete optimality system (93).

Remark 6.1. *Actually, the base functions of the space V_{h_s} and the space K_h can be taken differently. Often the control variable of a constrained control problem is less regular than the state variable. Due to the limited regularity of the optimal control in general, we normally use the piecewise constant space for approximation of the control, though linear continuous finite element spaces will be used to approximate the state and the co-state.*

Below we present a numerical example to illustrate our proposed Galekin formulation in Section 5 for stochastic control problem.

For simplicity in calculation, we take space domain $D = [-1, 1]$ and each stochastic domain Γ_i are $[-\sqrt{3}, \sqrt{3}]$ after finite dimensional representing of random fields. We assume each probability density function on Γ_i is uniform, i.e., $\rho_i(\xi_i) = \frac{1}{2\sqrt{3}}$, $i = 1, \cdots, N$. Thus, the joint probability density function $\rho(\xi)$ of random variable $\xi = (\xi_1, \xi_2, \cdots, \xi_N)$ is $\frac{1}{(2\sqrt{3})^N}$. In the following numerical example, we will do the same K-L expansion as [29] for random coefficient $a(x, \omega)$, i.e.

$$a_N(x,\xi) = \mathbb{E}a(x,\xi) + \sum_{n=1}^{N} \sqrt{\lambda_n}\phi_n(x)\xi_n,$$

where $(\lambda_n, \phi_n)_{1 \leqslant n \leqslant N}$ are eigenpairs of

$$\int_D e^{-|x_1-x_2|}\phi_n(x_1)dx_1 = \lambda_n \phi_n(x_2).$$

The first example we consider is the optimal control problem without any constraint on the control. We can see from Figure 1 and Table 1 that the values of $\mathcal{J}(y_h, u_h)$ are decreasing as the value of α becomes smaller.

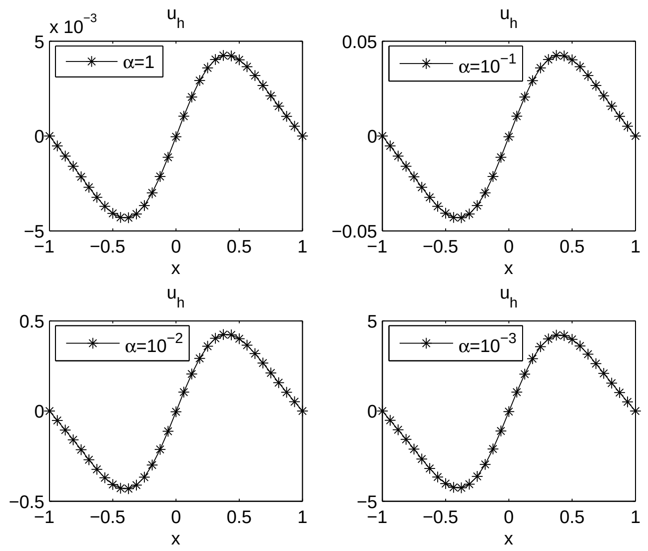

FIGURE 1. $N = 2, \mathbb{E}(a) = 29$.

Example 1 We consider the following model problem:

$$\min_{u\in K} \mathcal{J}(u) = \min_{u\in K}\Big(\frac{1}{2}\int_{-\sqrt{3}}^{\sqrt{3}} \frac{1}{(2\sqrt{3})^N}\int_{-1}^{1} |y-y_d|^2 dx d\xi + \frac{\alpha}{2}\int_{-1}^{1}|u|^2 dx\Big) \tag{102}$$

subject to

(103)

$$\begin{cases} -\nabla\cdot[a(x,\xi)\nabla y(x,\xi)] - \int_D (x-s)y(s,\xi)ds = u(x), & x\in[-1,\,1],\, \xi\in[-\sqrt{3},\sqrt{3}]^N,\\ y(x,\xi)=0, & x\in\{-1,\,1\},\ \xi\in[-\sqrt{3},\sqrt{3}]^N,\end{cases}$$

the target solution $y_d = \sin(2\pi x) + 3x(1-x^2)$.

TABLE 1. $N = 2, p = (1,1), \mathbb{E}(a) = 29$.

N	p	$\mathbb{E}(\lVert y_h - y_d\rVert^2)$	$\lVert u_h\rVert^2$	$\mathcal{J}(y_h,u_h)$	α
2	$(1,1)$	2.08113091531565	0.0000169265093185939	1.04057392091248	1
2	$(1,1)$	2.08082635323825	0.00169229096310293	1.04049779116728	10^{-1}
2	$(1,1)$	2.0777860549842	0.168869794671243	1.03973737646546	10^{-2}
2	$(1,1)$	2.04790687308469	16.5341986339144	1.0322205358593	10^{-3}

In the following three examples, we consider constrained optimal control with different types of constraints.

Example 2 We consider the following model problem:

$$\min_{u\in K} \mathcal{J}(u) = \min_{u\in K}\Big(\frac{1}{2}\int_{-\sqrt{3}}^{\sqrt{3}} \frac{1}{(2\sqrt{3})^N}\int_{-1}^{1} |y-y_d|^2 dx d\xi + \frac{1}{2}\int_{-1}^{1} |u|^2 dx\Big) \tag{104}$$

subject to

$$\begin{cases} -\nabla\cdot[0.01a(x,\xi)\nabla y(x,\xi)] - \int_D 0.5(x-s)y(s,\xi)ds = u(x), \\ \ x\in[-1,\,1],\ \xi\in[-\sqrt{3},\sqrt{3}]^N, \\ y(x,\xi)=0,\ x\in\{-1,\,1\},\ \xi\in[-\sqrt{3},\sqrt{3}]^N, \end{cases} \tag{105}$$

the target solution y_d=10$(\sin(2\pi x)+3x(1-x^2))$ and the deterministic control is constrained by the condition $u\geqslant 0, x\in[-1,\,1]$.

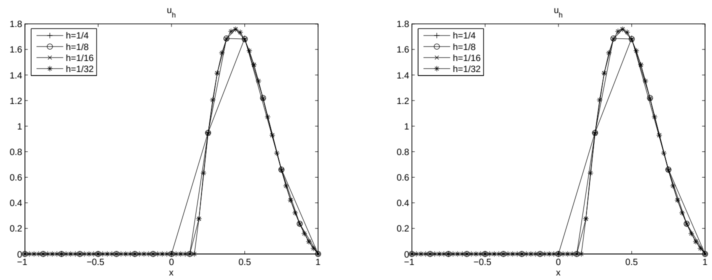

FIGURE 2. $N=2, \mathbb{E}(a)=29$, $p=(1,1)$(left), $p=(2,2)$(right).

TABLE 2. $N=2, p=(1,1), \mathbb{E}(a)=29$.

N	p	$\mathbb{E}(\Vert y_h-y_d\Vert^2)$	$\Vert u_h\Vert^2$	$\mathcal{J}(y_h,u_h)$	h
2	$(1,1)$	201.528597323722	0.924911032008952	101.226754177865	1/4
2	$(1,1)$	201.178840704619	1.02246463447778	101.100652669548	1/8
2	$(1,1)$	201.111044577624	1.04385831776585	101.077451447695	1/16
2	$(1,1)$	201.090356813128	1.0531131299507	101.071734971539	1/32

It follows from Figure 2 and Tables 2-3 that values of the objective function are decreasing as the mesh size decreases.

TABLE 3. $N=2, p=(2,2), \mathbb{E}(a)=29$.

N	p	$\mathbb{E}(\Vert y_h-y_d\Vert^2)$	$\Vert u_h\Vert^2$	$\mathcal{J}(y_h,u_h)$	h
2	$(2,2)$	201.528596545007	0.924910324740699	101.226753434873	1/4
2	$(2,2)$	201.178839689944	1.02246397880542	101.100651834375	1/8
2	$(2,2)$	201.111043610814	1.04385769516078	101.077450652987	1/16
2	$(2,2)$	201.090355806642	1.05311250192411	101.071734154283	1/32

Example 3 Consider the model problem (104) and (105), where $\alpha=1, \mu=0.03$, the target solution y_d=10$(\sin(2\pi x)+3x(1-x^2))$ and the deterministic control is constrained by the condition $-3.5\leqslant u\leqslant 3$.

It is clear from Figure 3 and Tables 4-5 that the approximated control is in the constraint, and values of the objective function decrease as the size of meshes does.

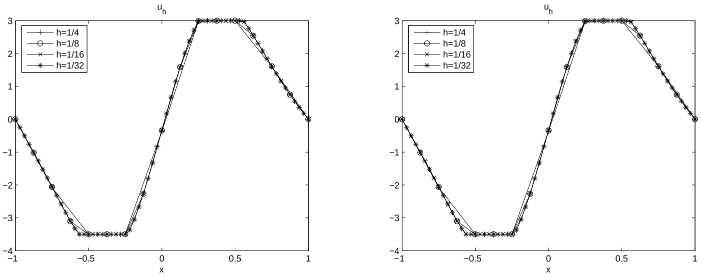

FIGURE 3. $N = 2, \mathbb{E}(a) = 29$, $p = (1, 1)$(left), $p = (2, 2)$(right)

TABLE 4. $N = 2, p = (1, 1), \mathbb{E}(a) = 29$.

N	p	$\mathbb{E}(\|y_h - y_d\|^2)$	$\|u_h\|^2$	$\mathcal{J}(y_h, u_h)$	h
2	$(1, 1)$	187.084122161325	11.0208385551139	99.0524803582197	1/4
2	$(1, 1)$	185.717118679463	11.6456477556407	98.6813832175521	1/8
2	$(1, 1)$	185.27576541775	11.886350006688	98.5810577122192	1/16
2	$(1, 1)$	185.188207511686	11.9295404646998	98.5588739881933	1/32

TABLE 5. $N = 2, p = (2, 2), \mathbb{E}(a) = 29$.

N	p	$\mathbb{E}(\|y_h - y_d\|^2)$	$\|u_h\|^2$	$\mathcal{J}(y_h, u_h)$	h
2	$(2, 2)$	187.084117222642	11.0208411177172	99.0524791701797	1/4
2	$(2, 2)$	185.717113110764	11.6456512722558	98.6813821915102	1/8
2	$(2, 2)$	185.275759842829	11.8863535896388	98.5810567162341	1/16
2	$(2, 2)$	185.188202170129	11.9295440354994	98.5588731028143	1/32

Example 4 Consider the model problem (104) and (105), the target solution y_d=$10(\sin(2\pi x) + 3x(1 - x^2))$ and the deterministic control is constrained by the condition $\int_{-1}^{1} u dx \geqslant 0$.

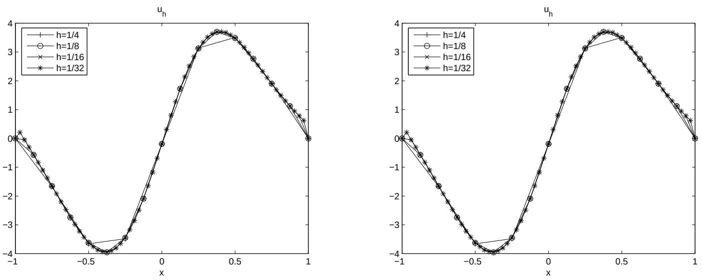

FIGURE 4. $N = 2, \mathbb{E}(a) = 29$, $p = (1, 1)$(left), $p = (2, 2)$(right).

TABLE 6. $N = 2, p = (1, 1), \mathbb{E}(a) = 29$.

N	p	$\mathbb{E}(\|y_h - y_d\|^2)$	$\|u_h\|^2$	$\mathcal{J}(y_h, u_h)$	h
2	$(1, 1)$	184.809790995933	12.0740992476723	98.441945121803	1/4
2	$(1, 1)$	182.983643748661	12.9607582106332	97.9722009796474	1/8
2	$(1, 1)$	182.49661209168	13.2041882360159	97.8504001638488	1/16
2	$(1, 1)$	182.372799177684	13.2704725910478	97.8216358843662	1/32

TABLE 7. $N = 2, p = (2, 2), \mathbb{E}(a) = 29$.

N	p	$\mathbb{E}(\|y_h - y_d\|^2)$	$\|u_h\|^2$	$\mathcal{J}(y_h, u_h)$	h
2	$(2,2)$	184.80978128723	12.0741071621808	98.4419442247058	1/4
2	$(2,2)$	182.983633162667	12.9607666885393	97.9721999256033	1/8
2	$(2,2)$	182.496601270413	13.2041968452073	97.8503990578106	1/16
2	$(2,2)$	182.372788296658	13.2704812248916	97.8216347607751	1/32

It follows from Figure 4 and Tables 6-7 that similar conclusions can be made for this type of control constraints.

References

[1] I. Babuska and P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, Comput. Methods Appl. Mech. Engrg., 191 (2002) 4093-4122.

[2] I. Babuska, K. Liu and R. Tempone, Solving stochastic partial differential equations based on the experimental data, Math. Models Methods Appl. Sci., 13 (2003) 415-444.

[3] I. Babuska, R. Tempone and G. E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM J. Numer. Anal., 42 (2004) 800-825.

[4] R. Becker and B.Vexler, Optimal control of the convection-diffusion equation using stabilized finite element methods, Numer. Math., 106 (2007) 349-367.

[5] S. C. Brenner and L. R. Scott, The mathematical theory of finite element methods, Second Edition, Springer, 2002.

[6] J. R. Cannon and Y. Lin, A priori L2 error estimates for Galerkin method for nonlinear parabolic integro-differential equations, Manuscript, 1987.

[7] P. G. Ciarlet, The finite element method for elliptic problems, In: Classics Appl.Math., vol. 40. SIAM, Philadelphia (2002).

[8] M. K. Deb, I. Babuska and J. T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Comput. Methods Appl. Mech. Engrg., 190 (2001) 6359-6372.

[9] I. Doltsinis, Inelastic deformation processes with random parametersmethods of analysis and design, Comput. Methods Appl. Mech. Engrg., 192 (2003) 2405-2423.

[10] L. Evans, Partial differential equations, Grad. Stud. Math. 19, AMS, Providence, RI, 1998.

[11] F. S. Falk, Approximation of a class of optimal control problems with order of convergence estimates. J. Math. Anal. Appl., 44 (1973) 28-47.

[12] A. Friedman and M. Shinbrot, Volterra integral equations in Banach space, Trans. Amer. Math. Soc., 126 (1967) 131-179.

[13] A. V. Fursikov, Optimal control of distributed systems, Theory and Applications, American Mathematical Society Providence, Rhode Island, 2000.

[14] R. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, Comput. Methods Appl. Mech. Engrg., 168 (1999) 19-34.

[15] R. Ghanem and S. Dham, Stochastic finite element analysis for multiphase flow in heterogeneous porous media, Trans. Porous Media, 32 (1998) 239-262.

[16] R. Ghanem and P. D. Spanos, Stochastic finite elements: a spectral approach, Springer-Verlag, Berlin, 1991.

[17] R. C. Grimmer and A. J. Pritchard, Analytic resolvent operators for integral equations in Banach space, J. Differential Equations, 50 (1983) 234-259.

[18] P. Grisvard, Elliptic problems in nonsmooth doamin, Longman Higher Education, 1986.

[19] 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.

[20] J. Haslinger and P. Neittaanmaki, Finite element approximation for optimal shape design, John Wiley and Sons, Chichester, 1989.

[21] M. L. Heard, An abstract parabolic Volterra integro-differential equation, SIAM J. Math. Anal., 13 (1982) 81-105.

[22] B. Hermann and N. N. Yan, Finite element methods for optimal control problems governed by integral equations and integro-differrential equations, Numer. Math., 101 (2005) 1-27.

[23] 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.

[24] A. Keese, Numerical solution of systems with stochastic uncertainties: a general purpose framework for stochastic Finite elements, Ph.D. thesis, Technical University Braunschweig, Braunschweig, Germany, 2004.
[25] A. Keese and H. G. Matthies, Parallel solution of stochastic PDEs, Proc. Appl. Math. Mech., 2 (2003) 485-486.
[26] M. Kleiber and T. D. Hien, The stochastic finite element method, John Wiley, Chichester, 1992.
[27] G.Knowles, Finite element approximation of parabolic time optimal control problems, SIAM J.Cont. Optim., 20 (1982) 414-427.
[28] A. Kufner, O. John and S. Fucik, Fuction Spaces, Nordhoff, Leiden, The Netherlands, 1977.
[29] H. C. Lee and J. Lee, A stochastic Galerkin method for stochastic control problems, Commum. Comput. Phys., 14 (2013) 77-106.
[30] M. N. LeRoux and V. Thome, Numerical solution of semilinear integro-differential equations of parabolic type with nonsmooth data, SIAM J. Numer. Anal., 26 (1989) 1291-130.
[31] R. Li, On Multi-Mesh h-Adaptive Algorithm, J. Sci. Comput., 24 (2005) 321-341.
[32] Y. Lin, V. Thome and L. Wahlbin, Ritz-Volterra projections to finite element spaces and applications to integro-differential and related equations, SIAM J. Numer. Anal., 28 (1991) 1047-1070.
[33] J. L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.
[34] J. L. Lions and E. Magenes, Non homogeneous boundary value problems and applications, Grandlehre B. 181, Springer-Verlag, 1972.
[35] W. B. Liu and D. Tiba, Error estimates for the finite element approximation of a class of nonlinear optimal control problems. J. Numer. Func. Optim., 22 (2001) 953-972.
[36] W. B. Liu and N. N. Yan, Adaptive finite element methods for optimal control governed by PDEs, Science Press, Beijing, 2008.
[37] M. Loève, Probability Theory, fourth ed., Springer, New York, 1997.
[38] A. Lorenzi and E. Sinestrari, An inverse problem in the theory of materials with memory, Nonlinear Anal. Theory Methods Appl., 12 (1988) 1317-1335.
[39] A. Lunardi and E. Sinestrari, C^{∞}-regularity of non-autonomous linear integro-differential equations of parabolic type, J. Diff. Equ., 63 (1986) 88-116.
[40] P. Malliavin, Stochastic analysis, Springer, Berlin, 1997.
[41] H. G. Matthies, C. E. Brenner, C. G. Bucher and C. GuedesSoares, Uncertainties in probabilistic numerical analysis of structures and solidsstochastic finite elements, Struct. Safety, 19 (1997) 283-336.
[42] H. G. Matthies and C. Bucher, Finite elements for stochastic media problems, Comput. Methods Appl. Mech. Engrg., 168 (1999) 3-17.
[43] P. Neittaanmaki and D. Tiba, Optimal control of nonlinear parabolic systems. Theory algorithms and applications, M.Dekker, New York, 1994.
[44] B. Øksendal, Stochastic differential equations, An introducation with application, 5th ed., Spring-Verlag, Berlin, 1998.
[45] M. Papadrakakis and V. Papadopoulos, Robust and efficient methods for stochastic finite element analysis using Monte Carlo simulation, Comput. Methods Appl. Mech. Engrg., 134 (1996) 325-340.
[46] O. Pironneau, Optimal shape design for elliptic systems, Springer-Verlag, Berlin, 1984.
[47] M. Renardy, W. J. Hrusa and J. A. Nohel, Mathematical Problems in Viscoelasticity, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 35, Longman Scientific and Technical, Harlow, Essex, 1987.
[48] E. Rosseel and G. N. Wells, Optimal control with stochastic PDE constrains and uncertain controls, Comput. Methods Appl. Mech. Engrg., 213-216 (2012) 152-167.
[49] G. I. Schuëller, A state-of-the-art report on computational stochastic mechanics, Prob. Engrg. Mech., 14 (1997) 197-321.
[50] C. Schwab and R. A. Todor, Sparse finite elements for elliptic problems with stochastic loading, Numer. Math., 95 (2003) 707-734.
[51] W. Shen, L. Ge and D. Yang, Finite element methods for optimal control problems governed by linear quasi-parabolic integro-differential equations, International Journal of Numerical Analysis and Modeling, 10 (2013) 536-550.
[52] W. Shen, L. Ge, D. Yang and W. Liu, A priori error estimates of finite element methods for linear parabolic integro-differential optimal control problems, Adv. Appl. Math. Mech., 6 (2014) 552-569.

[53] W. Shen, D. Yang and W. Liu, Optimal Control Problem Governed By A Linear Hyperbolic Integro-diffrential Equation and Its Finite Element Analysis, Bound. Value Prob., 2014:173, 1-29 (2014).

[54] I. H. Sloan and V. Thome, Time discretization of an integro-differential equation of parabolic type, SIAM J. Numer. Anal., 23 (1986) 1052-1061.

[55] V. Thome and N. Y. Zhang, Error estimates for semidiscrete finite element methods for parabolic integro-differential equations, Math. Comp., 53 (1989) 121-139.

[56] D. Tiba, Lectures on the optimal control of elliptic equations, University of Jyvaskyla Press, Finland (1995).

[57] R. Verfurth, A Review of a posteriori error estimation and adaptive mesh refinement, Wiley-Teubner, London, UK, 1996.

[58] N. Wiener, The homogeneous chaos, Amer. J. Math., 60 (1938) 897-936.

[59] D. Xiu and G. E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput., 24 (2002) 619-644.

[60] D. Xiu, D. Lucor, C. H. Su and G. E. Karniadakis, Stochastic modeling of flow-structure interactions using generalized polynomial chaos, ASME J. Fluid Engrg., 124 (2002) 51-69.

[61] E. G. Yanik and G. Fairweather, Finite element methods for parabolic and hyperbolic partial integro-differential equations, Nonlinear Anal., 12 (1988) 785-809.

School of Mathematic and Quantitative Economics, Shandong University of Finance and Economics. Jinan, 250014, P.R.China
E-mail: wanfangshen@gmail.com

School of Mathematics, Shandong University, Jinan, 250100, P.R. China
E-mail: tjsun@sdu.edu.cn

School of Mathematics, Shandong University, Jinan, 250100, P.R. China
E-mail: hgbxs@163.com

KBS, University of Kent, Canterbury, CT2 7NF, England
E-mail: W.B.Liu@kent.ac.uk