THE $h \times p$ FINITE ELEMENT METHOD FOR OPTIMAL CONTROL PROBLEMS CONSTRAINED BY STOCHASTIC ELLIPTIC PDES

HYUNG-CHUN LEE[1][†] AND JANGWOON LEE[2]

[1]DEPARTMENT OF MATHEMATICS, AJOU UNIVERSITY, SUWON, KOREA
E-mail address: hclee@ajou.ac.kr

[2]DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARY WASHINGTON, FREDERICKSBURG, VA 22401
E-mail address: llee3@umw.edu

ABSTRACT. This paper analyzes the $h \times p$ version of the finite element method for optimal control problems constrained by elliptic partial differential equations with random inputs. The main result is that the $h \times p$ error bound for the control problems subject to stochastic partial differential equations leads to an exponential rate of convergence with respect to p as for the corresponding direct problems. Numerical examples are used to confirm the theoretical results.

1. INTRODUCTION

The finite element method (FEM) has been used as a major tool to solve partial differential equations (PDEs), PDEs with random inputs or stochastic PDEs (SPDEs), and optimal control problems constrained by PDEs or SPDEs. There are three basic approaches to the FEM: the h version, the p version, and the h-p version of the FEM.

The h version of the FEM is the classical form of the FEM that has been used extensively for many years producing both theoretical and computational results. In the h version of the FEM, we fix the polynomial degree p, and then we reduce the mesh size h to obtain the desired accuracy. In contrast, in the p version of the FEM, we use a fixed mesh h and increase the polynomial degree p to get more accurate approximations. In the h-p version, we combine the h version and the p version approaches to get better results; i.e., we refine the mesh and increase the polynomial degree to have the numerical solutions converge faster to the exact solution.

In recent years, there has been an increasing interest in the h-p version of the Galerkin FEM (GFEM) including the discontinuous GFEM (DGFEM) [1, 2, 3, 4, 5] and the stochastic GFEM

(SGFEM) [6, 7, 8]. All these works are considering the h and/or p error bounds for only direct problems such as PDEs or SPDEs, not for control problems. It is well known that a control problem adds difficulties to the analysis of (stochastic) PDEs because of its flexible input data to control and its cost functional to be optimized; and hence, solving the control problem is a challenge. In our work we introduce the $h \times p$ version of the SGFEM for optimal control problems subject to stochastic elliptic problems. We then show for this $h \times p$ version approach that an exponential convergence rate with respect to p for the optimal control problems can be obtained.

The problem we consider is the optimization problem

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

constrained by the stochastic elliptic PDE under the Dirichlet boundary condition:

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

where $\mathbb{E}$ denotes expected value, D a convex bounded polygonal domain, ∂D its C^1 boundary, U a target solution to the constraint, β a positive constant that says the importance between two terms in (1.1), $a : D \times \Omega \to \mathbb{R}$ is a stochastic function with a bounded, continuous covariance function and a uniformly bounded, continuous first derivative, and $f \in L^2(D)$ a deterministic distributed control acting in the domain. For almost every $\omega \in \Omega$, with a flexible input f, we look for a solution u, stochastic function from $\overline{D} \times \Omega$ to $\mathbb{R}$ to optimize our functional (1.1). We would like to mention here that ∇ means differentiation with respect to $x \in D$ only.

To analyze our optimal control problem using the $h \times p$ version approach, we use the Karhunan-Loéve (KL) expansion [6, 7, 8, 9, 10, 11, 12, 13, 14, 15], transform stochastic PDEs to deterministic high dimensional PDEs, and present a priori error estimates of the $h \times p$ version of the SGFEM to the transformed model equation. After that, by using the method of Lagrange multipliers, we derive the optimality system of equations. Then we apply the theory of Brezzi-Rappaz-Raviart (BRR) [16, 17, 18, 19, 20] in uncoupling the optimality system, so that we can develop a priori error estimate that gives exponentially fast convergent results for the optimal solution of our optimal control problem.

Some remarks about the literature are in order. In 1980s, the p version was first studied in [21], and its theoretical results show that the p version gives results that are not worse then those obtained by the h version of the FEM when quasiuniform triangulations are used. In [22, 23, 24], the detailed error analysis of the h, p, and h-p versions of the FEM in a one dimensional setting was given. In [25], the exponential rate of the convergence of the h-p version is proved in a special geometric domain such as a square or a parallelogram. The work [26] analyzes the convergence of the h-p version of the FEM for elliptic problems with piecewise analytic data on curved domains. In [27], the author generalized the exponential rate of convergence of the h-p version of the FEM for elliptic equations of order $2m$.

In 1990s, it is shown in [28] that the auxiliary mapping technique in the frame of the p version of the FEM yields an exponential rate of convergence on domains with corners and

infinite domains. In [29], the fundamental theoretical ideas behind the p version and h-p version were discussed, and a benchmark comparison between the various versions was included. In [30], a new FEM method, the "partition of unity FEM", was presented that can be understood as a generalization of the h, p, and h-p versions of the FEM and more efficient than the usual FEM. In [31], the h-p version of the FEM was used to investigate the Galerkin FE solution to the Helmholtz equation.

In 2000s, the work [5] analyzes the h-p version of the DGFEM for the time discretization of parabolic equations to obtain exponential convergence results without severe restrictions on the space discretization. In [2], optimal convergence rates for the h-p version of the DGFEM to general advection diffusion-reaction problems were shown. In [3], the h-p version of the DGFEM was considered for the biharmonic equation to establish an a prior error estimate which is of optimal order with respect to the mesh h and nearly optimal with respect to the polynomial degree p. The work [1] uses the h-p version interior penalty DGFEM for second-order linear reaction-diffusion equations to obtain improved optimal error estimates. It is shown in [6] that the $h \times p$ version of the SGFEM method yields an exponential rate of convergence with respect to p for stochastic elliptic problems. In [4], the DGFEM to the biharmonic equation was used to obtain the h-p version bounds that are optimal with respect to the mesh size h and suboptimal with respect to the degree of the piecewise polynomial p.

There have been also papers by other authors published on the subject of optimal control with SPDE constraints (e.g., see recent three publications [32, 33, 34] and references therein). In [32], the authors proved the uniqueness of the optimal solution to the stochastic saddle problem after showing that it is equivalent to their optimality system. In [33], the computational solutions of optimal control problems constrained by SPDEs with uncertain controls were investigated, demonstrating the application of their methods via numerical examples. In the work [34], the authors examined the use of stochastic collocation for the numerical solution of optimal control problems subject to SPDEs, discussing generalized polynomial chaos thoroughly and presenting computational examples to show the performance of their method. Also, after finishing this paper the authors became aware of the work [35] including recent developments on the adaptive stochastic Galerkin FEM approches, which gives us some future research ideas. In the work [35], the authors developed adaptive refinement algorithms for SGFEM for countably-parametric, elliptic boundary value problems and proved the convergence of their adaptive algorithm.

The plan of the paper is as follows. In Section 2, we introduce stochastic function spaces and notations and then mention the uniqueness of the solution to our direct problems. In Section 3, we deal models with finite dimensional information and their $h \times p$ version error bounds on the solution of the models. In Section 4, we establish the $h \times p$ version error estimates to the discrete approximations of the solution to our optimal control problem. Finally, in Section 5, we give numerical examples of stochastic optimal control problems to verify our theoretical results.

2. Uniqueness of the solution

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is a set of outcomes, $\mathcal{F}$ is a σ-algebra of events, and $P : \mathcal{F} \to [0, 1]$ is a probability measure. Also, we use standard Sobolev space notation (see [36]). With this in mind, we define a stochastic Sobolev space as follows using strongly measurable functions:

$$L^2(\Omega; H_0^1(D)) = \{v : D \times \Omega \to \mathbb{R} \mid \|v\|_{L^2(\Omega; H^1(D))} < \infty\},$$

where

$$\|v\|_{L^2(\Omega; H_0^1(D))}^2 = \int_\Omega \|v\|_{H_0^1(D)}^2 \, dP = \mathbb{E}[\|v\|_{H_0^1(D)}^2].$$

Similarly, we can define $L^2(\Omega; L^2(D))$. Then for simplicity, we let $\mathcal{L}^2(D) = L^2(\Omega; L^2(D))$ and $\mathcal{H}_0^1(D) = L^2(\Omega; H_0^1(D))$.

We use the following notations to make our problem expressed easily:

$$b[u, v] = \mathbb{E}\left[\int_D a\nabla u \cdot \nabla v \, dx\right] \tag{2.1}$$

and

$$[u, v] = \mathbb{E}\left[\int_D uv \, dx\right]. \tag{2.2}$$

Using notations (2.1) and (2.2), we have the weak formulation of (1.2): seek $u \in \mathcal{H}_0^1(D)$ such that

$$b[u, v] = [f, v] \,\forall v \in \mathcal{H}_0^1(D). \tag{2.3}$$

In this paper, we assume that there are positive m and M such that

$$m \le a(x, \omega) \le M \quad \text{a.e. } (x, \omega) \in D \times \Omega. \tag{2.4}$$

For the condition (2.4), as a practical example, a could have a log normal distribution (see [37]).

Then from the Lax-Milgram lemma (see [38]), for $f \in L^2(D)$, we have a unique solution to (2.3).

3. The $h \times p$ version error bounds for high-dimensional models

3.1. **High-dimensional problems.** We use the KL expansions (e.g., see [7, 8, 39, 40]) to analyze our problem: If $a(x, \omega)$ is a stochastic function that has a continuous and bounded covariance function, it can be represented by

$$a(x, \omega) = \mathbb{E}[a(x, \omega)] + \sum_{n \ge 1} \sqrt{\lambda_n} \phi_n(x) X_n(\omega), \tag{3.1}$$

where the real random variables, $\{X_n\}$, are mutually uncorrelated, $\mathbb{E}[X_n X_m] = \delta_{nm}$, $\mathbb{E}[X_n] = 0$, and (λ_n, ϕ_n) are solutions to

$$\int_D C(x_1, x_2)\phi_n(x_1) \, dx_1 = \lambda_n \phi_n(x_2), \tag{3.2}$$

where $C(x_1, x_2) = \mathbb{E}[a(x_1, \omega)a(x_2, \omega)] - \mathbb{E}[a(x_1, \omega)]\mathbb{E}[a(x_2, \omega)]$. (3.1) is called the KL expansion of $a(x, \omega)$, and its convergence result can be found in [6].

In this paper, we use the truncated KL expansion of $a(x, \omega)$

$$a(x, \omega) = \mathbb{E}[a(x, \omega)] + \sum_{n=1}^{N} \sqrt{\lambda_n}\phi_n(x)X_n(\omega), \tag{3.3}$$

and we focus on our control problems to practical situations. To have the unique solution to the problem with (3.3), we assume that there exist $m, M > 0$ such that

$$m \leq a(x, \omega) \leq M \quad \text{a.e. } (x, \omega) \in D \times \Omega. \tag{3.4}$$

We also assume that each $X_n(\Omega) \equiv \Gamma_n \subset \mathbb{R}$ is a bounded interval for $n = 1, 2, \cdots, N$ and that each X_n has a density function $\rho_n : \Gamma_n \to \mathbb{R}^+$. We use the joint density $\rho(y)$ for any $y \in \Gamma = \prod_{n=1}^{N} \Gamma_n \subset \mathbb{R}^N$ of $(X_1, X_2, \cdots, X_N)$. Under these assumptions, the solution of (2.3) can be expressed by the finite number of random variables; i.e., $u(x, \omega) = u(x, X_1(\omega), X_2(\omega), \cdots, X_N(\omega))$; see e.g., [9, 6, 11]. Then we have the following high-dimensional deterministic equivalent weak formulation of (2.3) that we will focus on throughout the paper:

$$\int_{\Gamma} \rho(y) \int_{D} a(x, y)\nabla u(x, y) \cdot \nabla v(x, y) \, dxdy = \int_{\Gamma} \rho(y) \int_{D} f(x)v(x, y) \, dxdy. \tag{3.5}$$

The strong formulation of (3.5) is

$$\begin{aligned}
-\nabla \cdot [a(x, y)\nabla u(x, y)] &= f(x) \quad \forall (x, y) \in D \times \Gamma, \\
u(x, y) &= 0 \quad \forall (x, y) \in \partial D \times \Gamma
\end{aligned} \tag{3.6}$$

with its weak formulation: seek $u \in \mathcal{H}_0^1(D)$ such that for all $v \in \mathcal{H}_0^1(D)$,

$$b[u, v] = [f, v]; \tag{3.7}$$

Note that we have well-posedness of (3.6) because a is bounded and that by using the traditional finite element method, we can provide the solution to a SPDE from solving (3.6).

For our high-dimensional problem, we use the following space and notions:

$$\mathcal{H}_0^1(D) = L^2(\Gamma; H_0^1(D)) = \{v : D \times \Gamma \to \mathbb{R} \mid \|v\|_{L^2(\Gamma; H_0^1(D))} < \infty\},$$

$$b[u, v] = \int_{\Gamma} \rho \int_{D} a\nabla u \cdot \nabla v \, dxdy,$$

and

$$[u, v] = \int_{\Gamma} \rho \int_{D} uv \, dxdy.$$

3.2. Error bounds of high-dimensional problems. Let $X^h \subset H_0^1(D)$ and $G^h \subset L^2(D)$ be finite element spaces that consist of piecewise linear continuous functions. We assume that the following usual properties hold:

(i) for all $\phi \in H^2(D) \cap H_0^1(D)$, there exists $C > 0$ such that

$$\inf_{\phi^h \in X^h} \|\phi - \phi^h\|_{H_0^1(D)} \leq Ch\|\phi\|_{H^2(D)}; \tag{3.8}$$

(ii) for all $\phi \in H_0^1(D)$, there exists $C > 0$ such that

$$\inf_{\phi^h \in G^h} \|\phi - \phi^h\|_{L^2(D)} \leq Ch\|\phi\|_{H_0^1(D)}. \tag{3.9}$$

For the space Γ, we define the following finite element space with p_n that is the maximum degree of polynomial in a y_n-direction:

$$\mathcal{P}^\varphi = \mathcal{P}_1^{\varphi_1} \otimes \mathcal{P}_2^{\varphi_2} \otimes \cdots \otimes \mathcal{P}_N^{\varphi_N},$$

where $\varphi_n = 1/p_n$, $\varphi = (\varphi_1, \varphi_2, \cdots, \varphi_N)$, and

$$\mathcal{P}_n^{\varphi_n} = \{v : \Gamma_n \to \mathbb{R} : v \in span(1, y_n, y_n^2, \cdots, y_n^{1/\varphi_n})\}.$$

We now consider the following finite element weak formulation in $W = X^h \otimes \mathcal{P}^\varphi$: find $u^{h\varphi} \in W$ such that for all $v^{h\varphi} \in W$,

$$b[u^{h\varphi}, v^{h\varphi}] = [f, v^{h\varphi}]. \tag{3.10}$$

We also consider the following approximation property (e.g., see [6]).

Proposition 3.1. *Let* $u \in C^{p+1}(\Gamma; H^2(D) \cap H_0^1(D))$ *with* $p = (p_1, p_2, \cdots, p_N)$. *Then there exists a positive constant* C, *which is independent of* $h, \delta, N,$ *and* p, *such that*

$$\inf_{w \in W} \|u - w\|_{\mathcal{H}_0^1(D)} \leq C \left(h\|u\|_{\mathcal{H}^2(D)} + \delta^\gamma \sum_{j=1}^N \frac{\|\partial_{y_j}^{p_j+1} u\|_{\mathcal{H}_0^1(D)}}{(p_j+1)!} \right), \tag{3.11}$$

where $0 < \delta = \max_{1 \leq j \leq N} |\Gamma_j|/2 < 1$ *and* $\gamma = \min_{1 \leq j \leq N} \{p_j + 1\}$.

Then we have the following result follows from (3.11).

Proposition 3.2. *Let* $f(x) \in L^2(D)$, u *and* $u^{h\varphi}$ *be the solutions of* (3.7) *and* (3.10), *respectively. Then there exists* $C > 0$ *such that*

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} \leq C(h + \delta^\gamma)K\|f\|_{L^2(D)},$$

where $K = \sum_{j=1}^N \max\{1, \|\phi_j\|_{L^\infty(D)}^{p_j+1}\}$.

Remark 3.3. *With the same assumptions in Proposition 3.2, we have*

$$\|\mathbb{E}[u - u^{h\varphi}]\|_{H_0^1(D)} \leq C(h + \delta^\gamma)K, \tag{3.12}$$

where $K = \sum_{j=1}^N \max\{1, \sqrt{\lambda_j}\|\phi_j\|_{L^\infty(D)}^{p_j+1}\}$.

Remark 3.4. *For problems with $g(x, y) \in C^{p+1}(\Gamma; L^2(D))$, we have*

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} \leq C(h + \delta^\gamma) K \|g\|_{\mathcal{L}^2(D)}, \tag{3.13}$$

where $K = \sum_{j=1}^N \max\{1, \|\phi_j\|_{L^\infty(D)}^{p_j+1}, \sum_{k=1}^{p_j+1} \|\phi_j\|_{L^\infty(D)}^{p_j+1-k} \|\partial_{y_j}^k g\|_{\mathcal{L}^2(D)}\}$.

Note that Proposition 3.2 gives the optimal order of convergence with respect to δ (recall $\delta = \max\limits_{1 \leq j \leq N} |\Gamma_j|/2$), not with respect to p (recall $\gamma = \min\limits_{1 \leq j \leq N} \{p_j + 1\}$) unless we assume that each $|\Gamma_j| < 2$, $1 \leq j \leq N$. So, here we try to improve this situation in the following theorem without the need of assumptions on Γ_j.

Theorem 3.5. *Let $\epsilon \in (0, 1)$, $n \in \{1, 2, \cdots, N\}$, $f(x) \in L^2(D)$, and u and $u^{h\varphi}$ be the solutions of (3.7) and (3.10), respectively. Assume that there exists a constant $c > 0$, independent of N, such that*

$$\min_{x \in D}\{\mathbb{E}[a(x, y)] + \sum_{1 \leq j \leq N, j \neq n} \sqrt{\lambda_j}\phi_j(x)y_j\} - \sqrt{\lambda_n}\|\phi_n\|_{L^\infty(D)} \max_{y \in \Gamma_n} |y| \geq c > 0$$

for any $(y_1, y_2, \cdots, y_{n-1}, y_{n+1}, \cdots, y_N) \in \prod_{1 \leq j \leq N, j \neq n} \Gamma_j$. Then there exists $C > 0$ such that

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} \leq$$
$$C \left(h\|f\|_{L^2(D)} + \epsilon^{-1} \sum_{i=1}^N \sqrt{\pi|\Gamma|} \left(1 + (1 - r_i^2)^{-1/2}\mathcal{O}(\varphi_i^{1/3})\right) (r_i)^{1+1/\varphi_i} \right).$$

PROOF: Note that because the solution $u \in \mathcal{H}_0^1(D)$ of (3.5) is analytic with respect to $y \in \Gamma$ onto the space $H_0^1(D) \otimes \mathcal{P}^\varphi$ (e.g., see [6]), there exists a constant $C > 0$ such that

$$\min_{v \in H_0^1(D) \otimes \mathcal{P}^\varphi} \|u - v\|_{\mathcal{H}_0^1(D)} \leq \frac{C}{\epsilon c} \sum_{i=1}^N \sqrt{\pi|\Gamma|} \left(1 + (1 - r_i^2)^{-1/2}\mathcal{O}(\varphi_i^{1/3})\right) (r_i)^{1+1/\varphi_i},$$

where $0 < r_i \equiv (\sqrt{\sigma_i^2 - 1} + |\sigma_i|)^{-1} < 1$ and $\sigma_i < \dfrac{2c(\epsilon - 1)}{|\Gamma_i|\sqrt{\lambda_i}\|\phi_i\|_{L^\infty(D)}}$.

Note also that we have

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} \leq C \min_{v \in X^h \otimes \mathcal{P}^\varphi} \|u - v\|_{\mathcal{H}_0^1(D)}$$

$$\leq C \left(\min_{v \in X^h \otimes L^2(\Gamma)} \|u - v\|_{\mathcal{H}_0^1(D)} + \min_{v \in H_0^1(D) \otimes \mathcal{P}^\varphi} \|u - v\|_{\mathcal{H}_0^1(D)} \right)$$

$$\leq C \left(h\|u\|_{\mathcal{H}^2(D)} + \min_{v \in H_0^1(D) \otimes \mathcal{P}^\varphi} \|u - v\|_{\mathcal{H}_0^1(D)} \right)$$

for some $C > 0$.

Then the last two inequalities with H^2-regularity imply that

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} \le$$
$$C \left(h\|f\|_{L^2(D)} + \epsilon^{-1} \sum_{i=1}^N \sqrt{\pi|\Gamma|} \left(1 + (1 - r_i^2)^{-1/2} \mathcal{O}(\varphi_i^{1/3}) \right) (r_i)^{1+1/\varphi_i} \right)$$

for some positive constant C. □

Remark 3.6. *With the same assumptions in Theorem 3.5, there exist a constant $C > 0$ such that*

$$\|\mathbb{E}[u - u^{h\varphi}]\|_{H_0^1(D)} \le C(h + \epsilon^{-1} \sum_{i=1}^N (r_i)^{1+1/\varphi_i}), \tag{3.14}$$

where $0 < r_i < 1$ as in the proof of Theorem 3.5 above.

Remark 3.7. *For problems with $g(x, y) \in C^{p+1}(\Gamma; L^2(D))$, we have*

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} \le$$
$$C \left(h\|g\|_{\mathcal{L}^2(D)} + \epsilon^{-1} \sum_{i=1}^N \sqrt{\pi|\Gamma|} \left(1 + (1 - r_i^2)^{-1/2} \mathcal{O}(\varphi_i^{1/3}) \right) (r_i)^{1+1/\varphi_i} \right).$$

4. THE $h \times p$ VERSION ERROR BOUND FOR OPTIMAL CONTROL PROBLEMS

In this section, we derive the optimality system for our control problem constrained by elliptic PDE with random inputs and then analyze the control problem using the Brezzi-Rappaz-Raviart (BRR) theory obtaining its error bound. Throughout this section, we assume that $f \in L^2(D)$ for regularity of the solution.

4.1. **The optimality system.** We define the admissibility set as follows:

$$\mathcal{U}_{ad} = \{(u, f) \in \mathcal{H}_0^1(D) \times L^2(D) \text{ such that (2.3) satisfied and } \mathcal{J}(u, f) < \infty\}. \tag{4.1}$$

Then there exists an optimal solution $(\hat{u}, \hat{f}) \in \mathcal{U}_{ad}$ of $\mathcal{J}(u, f)$ such that $\mathcal{J}(\hat{u}, \hat{f}) \le \mathcal{J}(u, f)$ for all $(u, f) \in \mathcal{U}_{ad}$ satisfying $\|u - \hat{u}\|_{\mathcal{H}_0^1(D)} + \|f - \hat{f}\|_{L^2(D)} \le \epsilon$ for some $\epsilon > 0$.

Also, if $(u, f) \in \mathcal{H}_0^1(D) \times L^2(D)$ be an optimal solution of

$$\min_{(u,f)\in\mathcal{U}_{ad}} \mathcal{J}(u, f) \quad \text{subject to} \quad b[u, v] = [f, v] \; \forall v \in \mathcal{H}_0^1(D), \tag{4.2}$$

then there exists a Lagrange multiplier $\xi \in \mathcal{H}_0^1(D)$ such that

$$b[\xi, \zeta] = [u - U, \zeta] \quad \forall \zeta \in \mathcal{H}_0^1(D) \tag{4.3}$$

and

$$[\beta f + \xi, z] = 0 \quad \forall z \in L^2(D). \tag{4.4}$$

Remark 4.1. *The system formed by equations*

$$b[u, v] = [f, v] \; \forall v \in \mathcal{H}_0^1(D),$$
$$b[\xi, \zeta] = [u - U, \zeta] \quad \forall \zeta \in \mathcal{H}_0^1(D), \quad and \tag{4.5}$$
$$[\beta f + \xi, z] = 0 \quad \forall z \in L^2(D).$$

is called an optimality system. By solving this system, we can find the optimal solution of (4.2).

4.2. The Brezzi-Rappaz-Raviart theory.

Here for the sake of completeness, we will state the relevant result based on the BRR theory for our needs (for details, see [41, 42, 43]).

For Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, let $\mathcal{T} \in \mathcal{L}(\mathcal{Y}; \mathcal{X})$ and $\mathcal{G}$ be a C^2 mapping from $\mathcal{X}$ into $\mathcal{Y}$. Then we consider the following nonlinear problem: Seek $\psi \in \mathcal{X}$ such that

$$\psi + \mathcal{T}\mathcal{G}(\psi) = 0. \tag{4.6}$$

We assume that ψ is a regular solution of (4.6) and that there exists a Banach space $\mathcal{Z}$ continuously embedded in $\mathcal{Y}$ such that

$$\mathcal{G}_\psi(\psi) \in \mathcal{L}(\mathcal{X}; \mathcal{Z}) \quad \forall \psi \in \mathcal{X}, \tag{4.7}$$

where $\mathcal{G}_\psi$ is the Fréchet derivative of $\mathcal{G}$ with respect to ψ.

Let $\mathcal{X}^h$ be a finite dimensional subspace of $\mathcal{X}$ and $\mathcal{T}^h \in \mathcal{L}(\mathcal{Y}; \mathcal{X}^h)$ be an approximating operator. We assume the following properties for $\mathcal{T}^h$:

$$\lim_{h \to 0} \|(\mathcal{T}^h - \mathcal{T})\omega\|_\mathcal{X} = 0 \quad \forall \omega \in \mathcal{Y}; \tag{4.8}$$

$$\lim_{h \to 0} \|\mathcal{T}^h - \mathcal{T}\|_{\mathcal{L}(\mathcal{Z};\mathcal{X})} = 0. \tag{4.9}$$

Then we set the following approximate problem to the nonlinear problem above: Seek $\psi^h \in \mathcal{X}^h$ such that

$$\psi^h + \mathcal{T}^h\mathcal{G}(\psi^h) = 0. \tag{4.10}$$

We finally assume that $D^2\mathcal{G}$ is bounded on all bounded sets of $\mathcal{X}$, where $D^2\mathcal{G}$ represents any and all second Fréchet derivatives of $\mathcal{G}$.

Then with all assumptions above, there exists a neighborhood $\mathcal{O}$ of the origin in $\mathcal{X}$ and, for $h \leq h_0$ small enough, a unique $\psi^h \in \mathcal{X}^h$ such that ψ^h is a regular solution of (4.10). Moreover, there exists a constant $C > 0$, independent of h, such that

$$\|\psi^h - \psi\|_\mathcal{X} \leq C\|(\mathcal{T}^h - \mathcal{T})\mathcal{G}(\psi)\|_\mathcal{X}. \tag{4.11}$$

4.3. Error bound of optimal control problems.

In this section, we first fit our optimality system and its discrete approximation into the BRR framework such as (4.6) and (4.10) so that we can use the BRR theory. Then by proving all assumptions in Section 4.2, we get a error bound for our control problem constrained by PDEs with random inputs.

We set $\mathcal{X} = \mathcal{H}_0^1(D) \times L^2(D) \times \mathcal{H}_0^1(D)$ and $\mathcal{Y} = H^{-1}(D) \times \mathcal{H}^{-1}(D)$ and define the linear operator $\mathcal{T} \in \mathcal{L}(\mathcal{Y}; \mathcal{X})$ as follows:

$$(\tilde{u}, \tilde{f}, \tilde{\xi}) = \mathcal{T}(\tilde{r}, \tilde{\tau})$$

if and only if

$$b[\tilde{u}, v] = [\tilde{r}, v] \quad \forall v \in \mathcal{H}_0^1(D), \tag{4.12}$$

$$b[\tilde{\xi}, \zeta] = [\tilde{\tau}, \zeta] \quad \forall \zeta \in \mathcal{H}_0^1(D), \tag{4.13}$$

and

$$[\beta \tilde{f} + \tilde{\xi}, z] = 0 \quad \forall z \in L^2(D). \tag{4.14}$$

We also define $\mathcal{G} : \mathcal{X} \to \mathcal{Y}$ by

$$\mathcal{G}(\tilde{u}, \tilde{f}, \tilde{\xi}) = (-\tilde{f}, -\tilde{u} + U).$$

Then the optimality system (2.3), (4.3), and (4.4) can be written as

$$(u, f, \xi) + \mathcal{T}(\mathcal{G}(u, f, \xi)) = 0. \tag{4.15}$$

We now set $\mathcal{X}^{h\varphi} = W \times G^h \times W$, where $W = X^h \otimes \mathcal{P}^\varphi$ and define the discrete operator $\mathcal{T}^{h\varphi} \in \mathcal{L}(\mathcal{Y}; \mathcal{X}^{h\varphi})$ as follows:

$$(\tilde{u}^{h\varphi}, \tilde{f}^h, \tilde{\xi}^{h\varphi}) = \mathcal{T}^{h\varphi}(\tilde{r}, \tilde{\tau})$$

if and only if

$$b[\tilde{u}^{h\varphi}, v^{h\varphi}] = [\tilde{r}, v^{h\varphi}] \quad \forall v^{h\varphi} \in W, \tag{4.16}$$

$$b[\tilde{\xi}^{h\varphi}, \zeta^{h\varphi}] = [\tilde{\tau}, \zeta^{h\varphi}] \quad \forall \zeta^{h\varphi} \in W, \tag{4.17}$$

and

$$[\beta \tilde{f}^h + \tilde{\xi}^{h\varphi}, z^h] = 0 \quad \forall z^h \in G^h. \tag{4.18}$$

Then the discrete optimality system

$$\begin{aligned} b[u^{h\varphi}, v^{h\varphi}] &= [f^h, v^{h\varphi}] \quad \forall v^{h\varphi} \in W, \\ b[\xi^{h\varphi}, \zeta^{h\varphi}] &= [u^{h\varphi} - U, \zeta^{h\varphi}] \quad \forall \zeta^{h\varphi} \in W, \quad \text{and} \\ [\beta f^h + \xi^{h\varphi}, z^h] &= 0 \quad \forall z^h \in G^h \end{aligned} \tag{4.19}$$

can be written as

$$(u^{h\varphi}, f^h, \xi^{h\varphi}) + \mathcal{T}^{h\varphi}(\mathcal{G}(u^{h\varphi}, f, \xi^{h\varphi})) = 0.$$

We now proceed to verify all assumptions mentioned in Section 4.2. To do this, we define first a space $\mathcal{Z} = L^2(D) \times \mathcal{L}^2(D)$, which is continuously embedded into $\mathcal{Y} = H^{-1}(D) \times \mathcal{H}^{-1}(D)$.

We denote the Fréchet derivative of $\mathcal{G}$ with respect to (u, f, ξ) by $D\mathcal{G}(u, f, \xi)$. Then we obtain for $(u, f, \xi) \in \mathcal{X}$,

$$D\mathcal{G}(u, f, \xi) \cdot (\tilde{u}, \tilde{f}, \tilde{\xi}) = (-\tilde{f}, -\tilde{u}) \quad \forall (\tilde{u}, \tilde{f}, \tilde{\xi}) \in \mathcal{X}.$$

We now state the following propositions to have the error analysis for our optimal control problems constrained by PDEs with random inputs.

Proposition 4.2. $D\mathcal{G}(u, f, \xi) \in \mathcal{L}(\mathcal{X}; \mathcal{Z})$ *for all* $(u, f, \xi) \in \mathcal{X}$.

PROOF: Note that we have

$$\|DG(u, f, \xi) \cdot (\tilde{u}, \tilde{f}, \tilde{\xi})\|_{\mathcal{Z}} = \|\tilde{f}\|_{L^2(D)} + \|\tilde{u}\|_{\mathcal{L}^2(D)}$$
$$\leq \|\tilde{f}\|_{L^2(D)} + \|\tilde{u}\|_{\mathcal{H}_0^1(D)} + \|\tilde{\xi}\|_{\mathcal{H}_0^1(D)} < \infty. \qquad (4.20)$$

Then the result follows using (4.20). □

Proposition 4.3. $D^2 G$ *is bounded on all bounded sets of* $\mathcal{X}$.

PROOF: This can be shown from the fact that for any $(u, f, \xi) \in \mathcal{X}$,

$$D^2 G(u, f, \xi) \cdot (\tilde{u}, \tilde{f}, \tilde{\xi}) = (0, 0) \quad \forall (\tilde{u}, \tilde{f}, \tilde{\xi}) \in \mathcal{X}. \quad □$$

Proposition 4.4. *For any* $(\tilde{r}, \tilde{\tau}) \in \mathcal{Y}$, $\|(\mathcal{T} - \mathcal{T}^{h\varphi})(\tilde{r}, \tilde{\tau})\|_{\mathcal{X}} \to 0$ *as* $h, \varphi \to 0$.

PROOF: We have

$$\|(\mathcal{T} - \mathcal{T}^{h\varphi})(\tilde{r}, \tilde{\tau})\|_{\mathcal{X}} = \|(\tilde{u} - \tilde{u}^{h\varphi}, \tilde{f} - \tilde{f}^h, \tilde{\xi} - \tilde{\xi}^{h\varphi})\|_{\mathcal{X}} \qquad (4.21)$$
$$= \|\tilde{u} - \tilde{u}^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\tilde{f} - \tilde{f}^h\|_{L^2(D)}$$

First, we show $\|\tilde{u} - \tilde{u}^{h\varphi}\|_{\mathcal{H}_0^1(D)} \to 0$ as $h, \varphi \to 0$.

Let $\epsilon > 0$ and let $\tilde{r} \in H^{-1}(D)$. Then there exists a sequence of C^∞ functions $\{\tilde{r}_n\} \subset L^2(D)$ such that $\|\tilde{r}_n - \tilde{r}\|_{H^{-1}(D)} \to 0$ as $n \to 0$. That is, for some $\tilde{r}_{n_0} \in \{\tilde{r}_n\}$, we have $\|\tilde{r}_{n_0} - \tilde{r}\|_{H^{-1}(D)} < \epsilon$.

We now use our weak formulation (3.7) with $\tilde{r}_{n_0}$ to get the following inequality for some $C > 0$:

$$\|\tilde{u} - \tilde{u}_{n_0}\|_{\mathcal{H}_0^1(D)} < \epsilon C,$$

where $\tilde{u}_{n_0}$ is a weak solution.

Similarly, with (3.10), we get the following result: there is a constant $C > 0$ such that

$$\|\tilde{u}^{h\varphi} - \tilde{u}_{n_0}^{h\varphi}\|_{\mathcal{H}_0^1(D)} < \epsilon C,$$

where $\tilde{u}_{n_0}^{h\varphi}$ is a weak solution.

These two inequalities imply that

$$\|\tilde{u} - \tilde{u}^{h\varphi}\|_{\mathcal{H}_0^1(D)} \leq \|\tilde{u} - \tilde{u}_{n_0}\|_{\mathcal{H}_0^1(D)} + \|\tilde{u}_{n_0} - \tilde{u}_{n0}^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\tilde{u}_{n_0}^{h\varphi} - \tilde{u}^{h\varphi}\|_{\mathcal{H}_0^1(D)}$$
$$\leq \|\tilde{u}_{n_0} - \tilde{u}_{n_0}^{h\varphi}\|_{\mathcal{H}_0^1(D)}$$

since ϵ is arbitrary. Then $\|\tilde{u} - \tilde{u}^{h\varphi}\|_{\mathcal{H}_0^1(D)} \to 0$ as $h, \varphi \to 0$ from Theorem 3.5.

Similarly, $\|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{H}_0^1(D)} \to 0$ as $h, \varphi \to 0$.

Now we want to show that $\|\tilde{f} - \tilde{f}^h\|_{L^2(D)} \to 0$ as $h \to 0$. Note from using (4.14) and (4.18), we obtain

$$[\tilde{f} - \tilde{f}^h, \tilde{f} - \tilde{f}^h] = [\tilde{f} - \tilde{f}^h, \tilde{f} - g^h] + 1/\beta[\tilde{\xi} - \tilde{\xi}^{h\varphi}, \tilde{f}^h - \tilde{f} + \tilde{f} - g^h].$$

Then by the Hölder, Cauchy, and Young inequalities, we have

$$
\begin{aligned}
\|\tilde{f} - \tilde{f}^h\|_{\mathcal{L}^2(D)}^2 \leq\ & \|\tilde{f} - \tilde{f}^h\|_{\mathcal{L}^2(D)}\|\tilde{f} - g^h\|_{\mathcal{L}^2(D)} \\
& + 1/\beta\|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{L}^2(D)}\|\tilde{f}^h - \tilde{f}\|_{\mathcal{L}^2(D)} \\
& + 1/\beta\|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{L}^2(D)}\|\tilde{f} - g^h\|_{\mathcal{L}^2(D)} \\
\leq\ & 1/4\|\tilde{f} - \tilde{f}^h\|_{\mathcal{L}^2(D)}^2 + \|\tilde{f} - g^h\|_{\mathcal{L}^2(D)}^2 \\
& + 1/\beta\|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{L}^2(D)}^2 + 1/4\|\tilde{f}^h - \tilde{f}\|_{\mathcal{L}^2(D)}^2 \\
& + 1/(2\beta)\|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{L}^2(D)}^2 + 1/2\|\tilde{f} - g^h\|_{\mathcal{L}^2(D)}^2.
\end{aligned}
\tag{4.22}
$$

The Inequality (4.22) implies that

$$
\|\tilde{f} - \tilde{f}^h\|_{\mathcal{L}^2(D)}^2 \leq C(\|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{L}^2(D)}^2 + \|\tilde{f} - g^h\|_{\mathcal{L}^2(D)}^2)
$$

for some $C > 0$. Here we note that from (4.14) and the Hölder inequality, we have $\int_D |\nabla \tilde{f}|^2\, dx < \infty$. With choosing $g^h = P^h(\tilde{f})$, where P^h is the standard L^2-projection operator from $L^2(D)$ to G^h, we have $\|\tilde{f} - g^h\|_{\mathcal{L}^2(D)}^2 \to 0$ as $h \to 0$. Thus, from the previous result above, as $h, \varphi \to 0$, we have

$$
\|\tilde{f} - \tilde{f}^h\|_{\mathcal{L}^2(D)}^2 \to 0.
$$

Therefore, with all arguments mentioned above, we have the following result:

$$
\|(\mathcal{T} - \mathcal{T}^{h\varphi})(\tilde{r}, \tilde{\tau})\|_{\mathcal{X}} = \|\tilde{u} - \tilde{u}^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\tilde{\xi} - \tilde{\xi}^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\tilde{f} - \tilde{f}^h\|_{L^2(D)} \to 0
$$

as $h \to 0$ and $\varphi \to 0$. □

Proposition 4.5. $\|\mathcal{T} - \mathcal{T}^{h\varphi}\|_{\mathcal{L}(\mathcal{Z}, \mathcal{X})} \to 0$ as $h, \varphi \to 0$.

PROOF: We note that $H^1(D)$ is compactly embedded in $L^2(D)$ if $D \subset \mathbb{R}^d$ is a bounded domain with C^1 boundary by Morrey's inequality and Arzela-Ascoli Compactness theorem with $d = 1$, by Rellich's theorem with $d = 2$, and by Rellich-Kondrachov Compactness theorem with $d = 3$. Thus, by compact embedding results, $\mathcal{Z} \subset \mathcal{Y}$ is compact. Thus, the proof of this proposition follows from the result in Proposition 4.4. □

Proposition 4.6. *A solution of* (4.15) *is regular.*

PROOF: A proof follows from the linearity and well-posedness of (4.12), (4.13), and (4.14). □

Through Propositions 4.2 - 4.6 we have verified all of the assumptions in Section 4.2. Thus, we obtain the following exponential convergence with respect to the polynomial degree $p = 1/\varphi$, which is our main result.

Theorem 4.7. *Assume that* $U \in \mathcal{H}_0^1(D)$. *Let* $(u, f, \xi) \in \mathcal{H}_0^1(D) \times L^2(D) \times \mathcal{H}_0^1(D)$ *be the solution of the optimality system* (4.5), $(u^{h\varphi}, f^h, \xi^{h\varphi}) \in W \times G^h \times W$ *be the solution of the*

discrete optimality system (4.19), and $\epsilon \in (0,1)$. Then we have

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\xi - \xi^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|f - f^h\|_{L^2(D)} \to 0 \text{ as } h, \varphi \to 0.$$

Moreover, there exists $C > 0$ such that

$$\|u - u^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|\xi - \xi^{h\varphi}\|_{\mathcal{H}_0^1(D)} + \|f - f^h\|_{L^2(D)} \leq \tag{4.23}$$

$$C \left(h[\|f\|_{L^2(D)} + \|u - U\|_{\mathcal{H}_0^1(D)}] + \frac{1}{\epsilon} \sum_{i=1}^N \sqrt{\pi |\Gamma|} \left[1 + \frac{\mathcal{O}(\varphi_i^{1/3})}{(1 - r_i^2)^{1/2}} \right] (r_i)^{1+1/\varphi_i} \right),$$

where $0 < r_i \equiv (\sqrt{\sigma_i^2 - 1} + |\sigma_i|)^{-1} < 1$ and $\sigma_i < \dfrac{2c(\epsilon - 1)}{|\Gamma_i|\sqrt{\lambda_i}\|\phi_i\|_{L^\infty(D)}}$ with a constant $c > 0$ and eigenpairs (λ_i, ϕ_i) in (3.3).

5. NUMERICAL COMPUTATION OF OPTIMAL CONTROL PROBLEMS

In this section, we verify our theoretical results using some numerical examples; i.e., we report on some numerical experiments to show the exponential convergence results with respect to the polynomial degree $p = (p_1, p_2, \cdots, p_N)$, where p_n is the maximum degree of polynomials in a y_n-direction.

5.1. Numerical Setting.
In our numerical experiments, we use that our deterministic domain D is $[-1, 1]$ and each stochastic domain Γ_n is $[-\sqrt{3}, \sqrt{3}]$. Also we suppose that we have a constant density function. Then from the assumptions about X_n in the KL expansion, we obtain the joint probability density function ρ of $(X_1, X_2, \cdots, X_N)$ in our numerical experiments is $(2\sqrt{3})^{-N}$.

Here we use a smooth covariance function $C(x_1, x_2) = e^{-|x_1 - x_2|/\alpha}$, where $\alpha > 0$ is the correlation length. Then

$$\int_D C(x_1, x_2)\phi_n(x_1)\, dx_1 = \lambda_n \phi_n(x_2).$$

gives the following eigenfunctions $\{\phi_n\}$ and eigenvalues $\{\lambda_n\}$:

$$\phi_n(x) = \cos(v_n x)\left(\sqrt{1 + (2v_n)^{-1}\sin(2v_n)}\right)^{-1/2} \quad \text{if n is odd,}$$

$$\phi_n(x) = \sin(w_n x)\left(\sqrt{1 - (2w_n)^{-1}\sin(2w_n)}\right)^{-1/2} \quad \text{if n is even,}$$

$$\lambda_n = 2\alpha(1 + \alpha^2 v_n^2)^{-1} \quad \text{if n is odd, and}$$

$$\lambda_n = 2\alpha(1 + \alpha^2 w_n^2)^{-1} \quad \text{if n is even,}$$

where v_n is a solution of $1 - \alpha v \tan(v) = 0$ and w_n is a solution of $\alpha w + \tan(w) = 0$. Note that with the correlation length $\alpha = 1$, our eigenvalues $\{\lambda_n\}$ decay quickly enough as n increases as we can see in Figure 5.1. Also, as we may guess from $\{\lambda_n\}$ above and from the literature (e.g., see [39, 44]), we see that the eigenvalue decay rate gets smaller when the correlation length gets smaller; for example, when $\alpha = 0.01$, the eigenvalue decay is hardly visible. Thus,
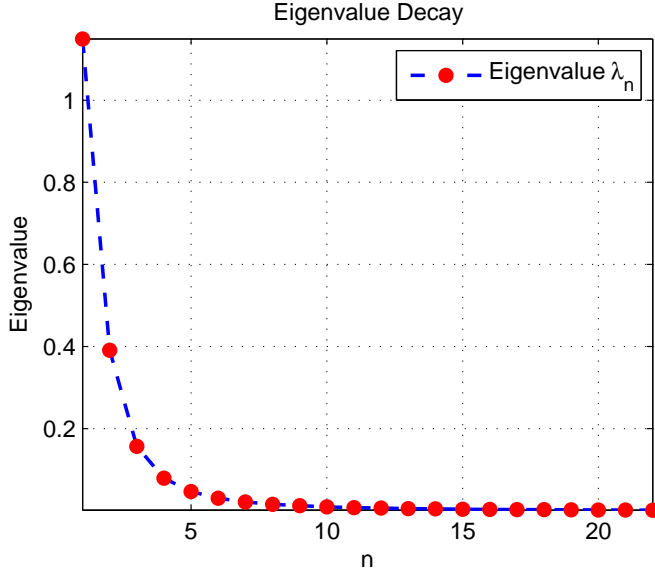
FIGURE 5.1. Eigenvalue Decay with $e^{-|x_1-x_2|}$

in this paper, we use $C(x_1, x_2) = e^{-|x_1-x_2|}$, which gives a reasonable eigenvalue decay rate so that we can use the first few terms of our KL expansion in numerical computation.

To confirm our theoretical results, we consider a model problem with the target solution $U = \sin(\pi x) + \sin(2\pi x)$: Find the solution of

$$-(a(x,y)u'(x,y))' = f(x) \quad \forall(x,y) \in (-1,1) \times \prod_{n=1}^{N}(-\sqrt{3}, \sqrt{3}), \qquad (5.1)$$

$$u(x,y) = 0 \quad \forall(x,y) \in \{-1,1\} \times \prod_{n=1}^{N}(-\sqrt{3}, \sqrt{3}),$$

where $a(x,y) = \mathbb{E}[a(x,y)] + \sum_{n=1}^{N} \sqrt{\lambda_n}\phi_n(x)y_n$ by controlling $f(x)$ to minimize

$$\mathcal{J}(u,f) = \frac{1}{2}\int_{-\sqrt{3}}^{\sqrt{3}} \frac{1}{(2\sqrt{3})^N} \int_{-1}^{1} |u - U|^2 \, dxdy + \frac{\beta}{2}\int_{-1}^{1} |f|^2 \, dx. \qquad (5.2)$$

5.2. **Numerical Results.** In this section, we provide the graphs of discrete optimal solutions together with the target solution; we give the tables of the values of relative errors for the optimal solutions with various polynomial degrees and their corresponding numbers of degrees of freedom of the discretization with respect to one-, two-, and three-dimensional stochastic domains; we finally present the figures of exponential convergence results with respect to $p = (p_1, p_2, \cdots, p_N)$ in three different dimensional spaces.

In tables and figures for computational results, for simplicity, we use $DP = \sum_{n=1}^{N} p_n$, where p_n is the maximum degree of polynomials in a y_n-direction and DOF as the number

$DP(DOF)$	Relative Error for u	Relative Error for ξ	Relative Error for f
1(2)	1.198032987982e-01	1.787953125199e-01	1.720684070324e-01
3(4)	9.562740236029e-03	1.524223854302e-02	1.341233756062e-02
5(6)	5.570898785792e-04	1.057333186124e-03	8.615989678717e-04
7(8)	2.795461789913e-05	6.309221819675e-05	4.901139558725e-05
9(10)	1.286833647831e-06	3.391536066446e-06	2.567358708683e-06
11(12)	4.560494644988e-08	1.362462543327e-07	1.019093421438e-07

TABLE 5.1. Dimension $= 1$, $\mathbb{E}[a(x,y)] = 2$

$DP(DOF)$	Relative Error for u	Relative Error for ξ	Relative Error for f
1(2)	9.401289567930e-02	1.987307188342e-01	1.752373609462e-01
3(6)	1.021973093216e-02	2.469886703133e-02	1.951724713049e-02
5(12)	1.005212637175e-03	3.036986195881e-03	2.380145213483e-03
7(20)	9.899411004937e-05	3.609536395869e-04	2.891208795498e-04
9(30)	9.451679011477e-06	3.927211535465e-05	3.149626856263e-05
11(42)	3.335005157516e-07	2.710535612892e-06	2.420352397715e-06

TABLE 5.2. Dimension $= 2$, $\mathbb{E}[a(x,y)] = 3$

$DP(DOF)$	Relative Error for u	Relative Error for ξ	Relative Error for f
1(2)	1.213041981628e-01	2.365082151581e-01	2.067441516861e-01
3(8)	3.705922498228e-02	6.324595325181e-02	5.759006750746e-02
5(18)	7.664352606772e-03	2.486450915935e-02	2.261773242123e-02
7(36)	1.714928430140e-03	8.617173569856e-03	9.261673667937e-03
9(64)	7.369214036585e-04	2.992784382786e-03	3.180578082762e-03
11(100)	1.650674990773e-05	1.329491864220e-04	1.672158841553e-04

TABLE 5.3. Dimension $= 3$, $\mathbb{E}[a(x,y)] = 3$

of degrees of freedom of the discretization with respect to the random parameter space. For instance, if we use $p = (p_1, p_2) = (5, 4)$, then $DP = 9$ and $DOF = 30$.

We in our numerical examples focus on the convergence of the discrete optimal solutions, $u^{h\varphi}$, $\xi^{h\varphi}$, and f^h, with respect to the polynomial degree p in terms of the relative error norms. For example, for the state solution u and the Lagrange multiplier ξ, we use the H_0^1 norm, and for our control f, we use the L^2 norm. Also, to satisfy the coercivity condition for our coefficient $a(x,y)$, we use appropriate values of $\mathbb{E}[a(x,y)]$ in each numerical experiment.

First, for the discrete optimal solutions' graphs, we run our programs by increasing the polynomial degree p with fixed step size on fixed spatial and stochastic domains as mentioned above. We then provide solution graphs together with the given target for only two dimensional

problem (we would like to mention that we have similar results in other dimensional cases as well). The graphs of the expectations of our optimal solutions $\mathbb{E}[u^{h\varphi}]$, $\mathbb{E}[\xi^{h\varphi}]$, and f^h with our target solution $U = \sin(\pi x) + \sin(2\pi x)$ for different values of $DP = 2, 4, 6, 8$, and 10 are shown in Figure 5.2.
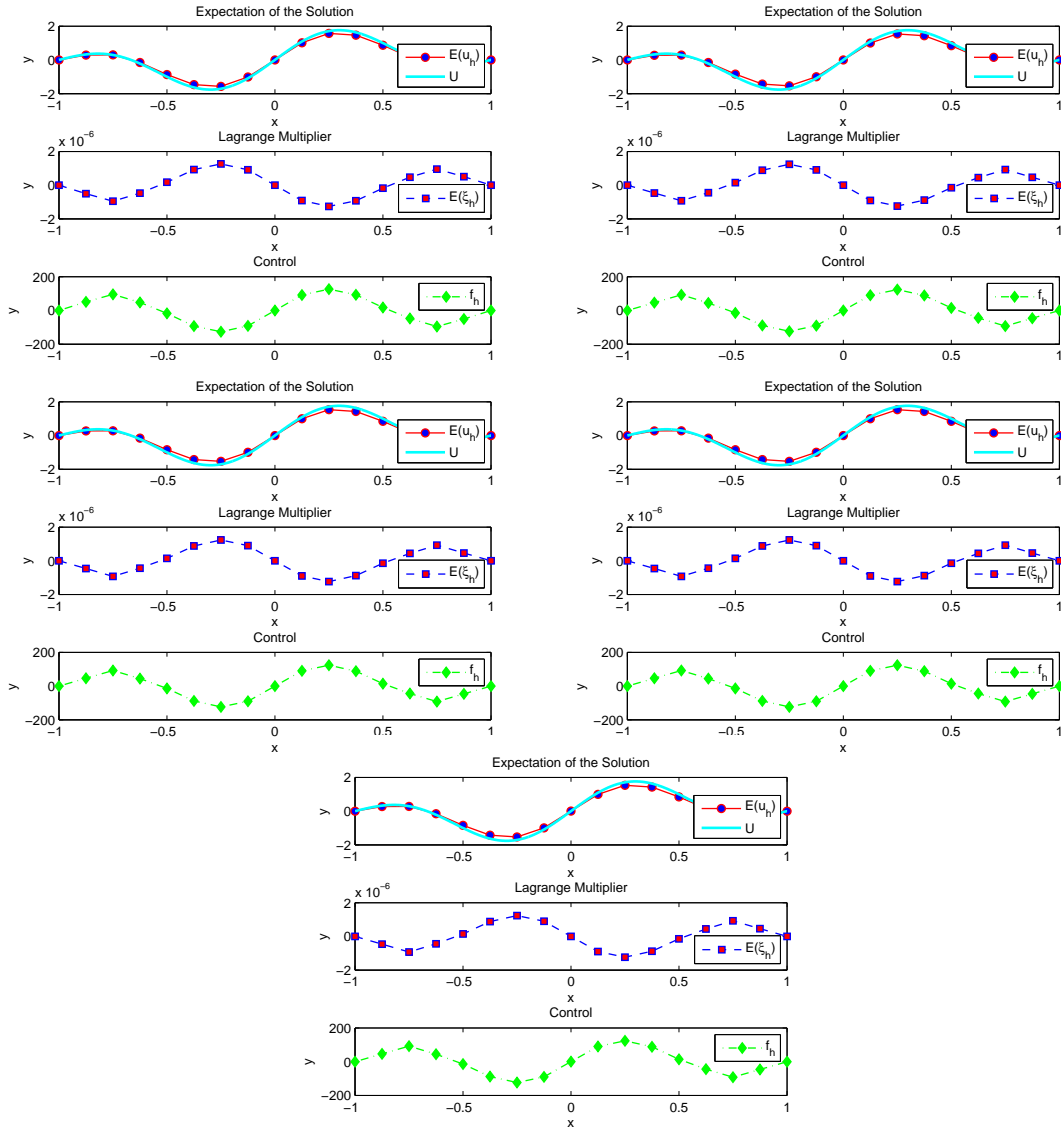

FIGURE 5.2. Dimension $= 2$, $\mathbb{E}[a(x, y)] = 3$, $U = \sin(\pi x) + \sin(2\pi x)$, $DP = 2$ (Top left), $DP = 4$ (Top right), $DP = 6$ (Middle left), $DP = 8$ (Middle right), $DP = 10$ (Bottom)

In addition, for each dimensional problem, we present figures of $\mathbb{E}[u^{h\varphi}]$ together with U, $\mathbb{E}[\xi^{h\varphi}]$, and f^h when DP is equal to 12 (see Figure 5.3). As we can see from Figures 5.2 and 5.3, in all different cases, we have the values of $\mathbb{E}[u^{h\varphi}]$ closer to the target U as desired.

Second, as we may expect from our theoretical convergence results, when we solve the discrete optimal control problems in one-, two-, and three-dimensional random parameter spaces, the relative errors for the optimal solutions get much smaller as the polynomial degree increases (see Tables 5.1, 5.2, and 5.3). Here we get errors a little bit more as we increase the dimension of the random parameter space, but these are reasonable results because the upper bound of each error include more and more terms as the dimension gets bigger and because the results may depend on the coercivity condition of $\mathbb{E}[a(x,y)]$.
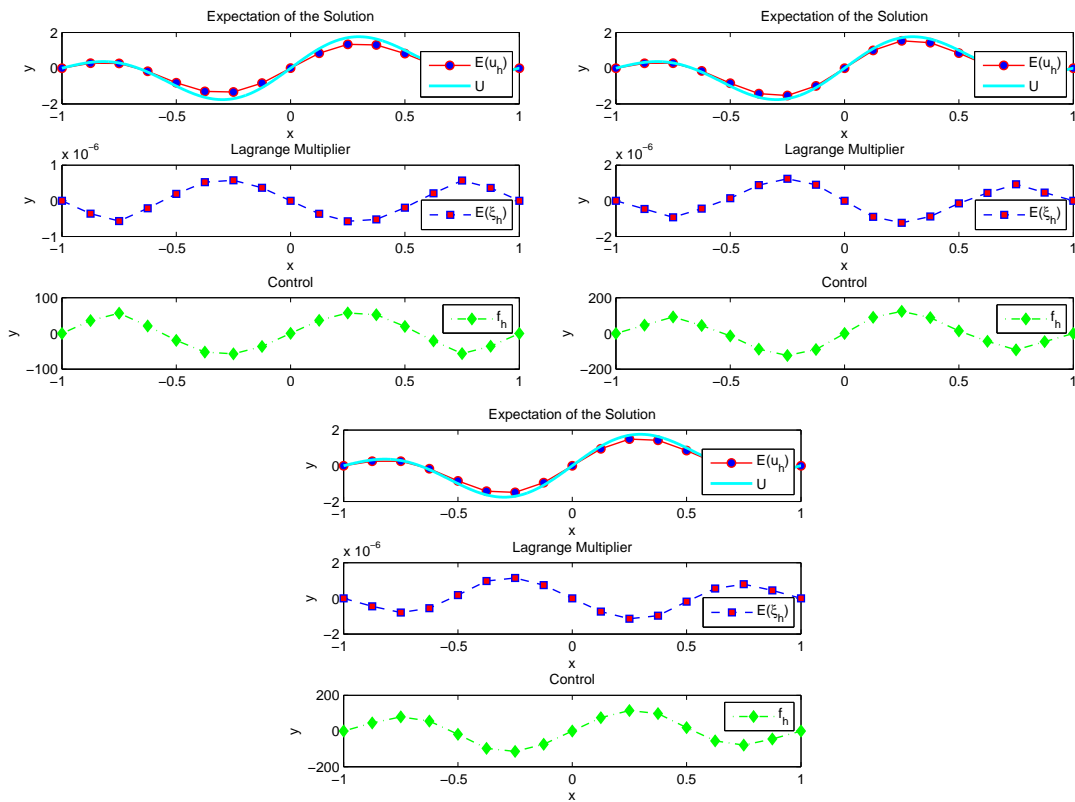

FIGURE 5.3. $U = \sin(\pi x) + \sin(2\pi x)$, $DP = 12$, $Dimension = 1$ (Top left), $Dimension = 2$ (Top right), $Dimension = 3$ (Bottom)

Finally, as shown in Figure 5.4, based on the values in tables, we obtain exponential rates of convergence for the discrete optimal solutions with respect to the polynomial degree; these numerical results confirm our theoretical convergence rates of the optimal solutions.
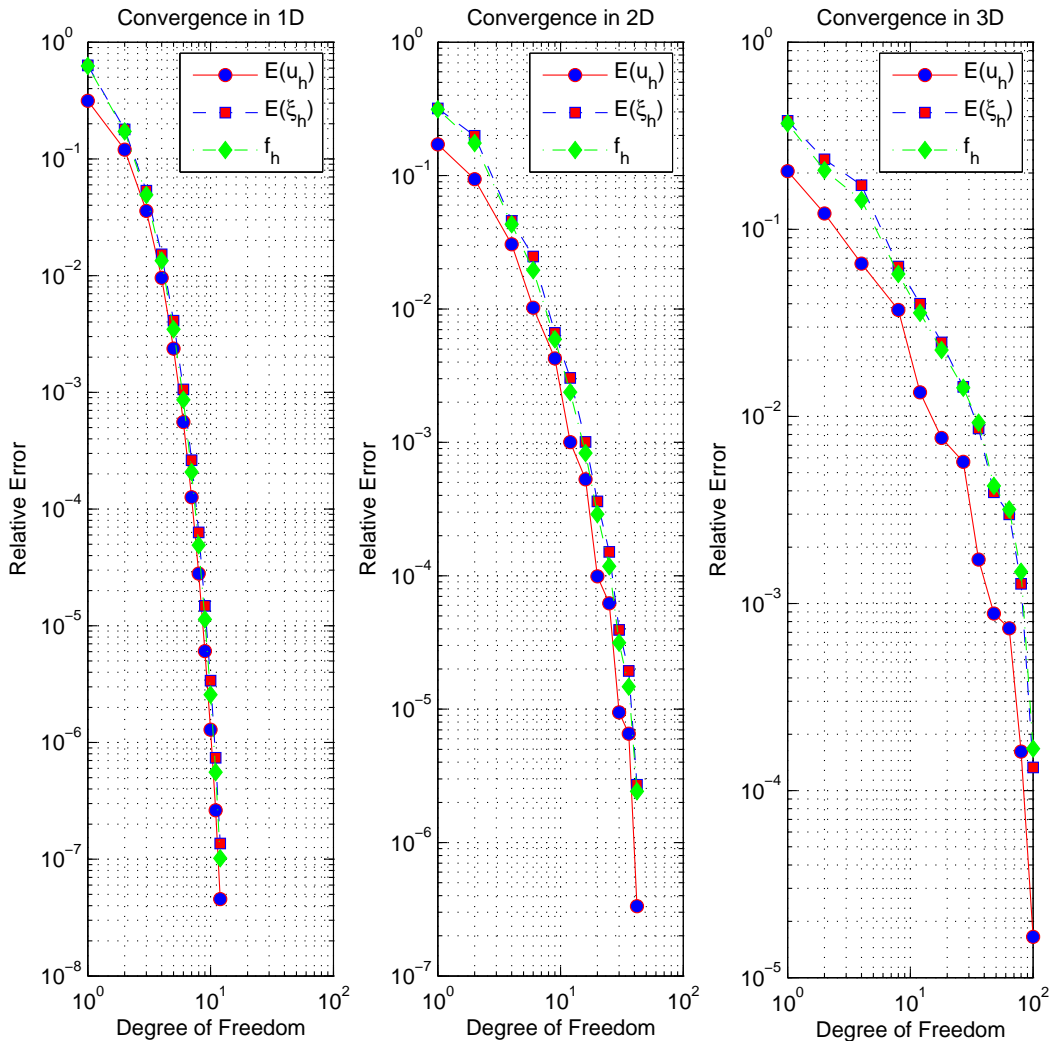
FIGURE 5.4. Convergence of Optimal Solutions

6. CONCLUSION REMARKS

Uncertainty can be found everywhere and we cannot just avoid it. In fact, many physical, biological, chemical, social, economic, and financial systems always involve some types of uncertainties. For example, we may consider media properties in oil reservoirs, which are very costly to obtain by measurement. Also, we may think of rainfall amounts that are unpredictable. From these examples, we see that data available are not always complete in modeling some phenomena. As a result, mathematical model equations describing theses phenomena should take uncertainty into account.

In our model problem (constraint equation or stochastic elliptic PDE), we introduced an element of randomness into incomplete input data to describe some degree of uncertainty of measurement. Then we analyzed the stochastic PDE constrained optimal control problem by reformulating in terms of a high dimensional parametric, deterministic problem. Although we used a stochastic Galerkin finite element method to solve the problem numerically, one could use a stochastic collocation method and/or a proper sparse grid discretization in a high dimensional space to get better computational results. In addition we could model the Navier-Stokes equation involving uncertainty and use our idea to analyze the equation describing the motion of fluids more accurately and giving better predictions. These problems will be addressed in our future papers.

References

[1] E. H. Georgoulis and E. Suli, *Optimal error estimates for the hp-version interior penalty discontinous Galerkin finite element method*, IMA J. Numer. Anal., **25** (2005), 205–220.

[2] P. Houston, C. Schwab, and E. Suli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., **39** (2002), 2133–2163.

[3] I. Mozolevski and E. Suli, *A priori error analysis for the hp-version of the discontinuous Galerkin finite element method for the biharmonic equation*, Comput. Meth. Appl. Math., **3** (2003), 1–12.

[4] I. Mozolevski, E. Suli, and P. R. Bosing, *hp-version a priori error analysis of interior penalty discontinuous Galerkin finite element approximations to the biharmonic equation*, J. Sci. Comput., **30** (2007), 465–491.

[5] D. Schotzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinous Galerkin finite element method*, SIAM J. Numer. Anal., **38** (2000), 837–875.

[6] 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.

[7] I. Babuska, R. Tempone, and G. E. Zouraris, *Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation*, Comput. Method Appl. Mech. Engrg., **194** (2005), 1251–1294.

[8] P. Frauenfelder, C. Schwab, and R. A. Todor, *Finite elements for elliptic problems with stochastic coefficients*, Comput. Method Appl. Mech. Engrg., **194** (2005), 205–228.

[9] I. Babuska and P. Chatzipantelidis, *On solving elliptic stochastic partial differential equations*, Comput. Method Appl. Mech. Engrg., **191** (2002), 4093–4122.

[10] 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.

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

[12] 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.

[13] H.-C. Lee and J. Lee, *A Stochastic Galerkin Method for Stochastic Control Problems*, Commun. Comput. Phys., **14** (2013), 77–106.

[14] 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. Vol., **384** (2011), 87–103.

[15] C. Schwab and R. A. Todor, *Sparse finite elements for elliptic problems with stochastic loading*, Numer. Math., (2003), 707–734.

[16] K. Chrysafinos, *Moving mesh finite element methods for an optimal control problem for the advection-diffusion equation*, J. Sci. Comput., **25** (2005), 401–421.

[17] M. D. Gunzburger, L. S. Hou, and T. Svobodny, *Analysis and finite element approximation of optimal control problems for the stationary Navier–Stokes equations with distributed and Neumann controls*, Math. Comp. **57** (1991), 123–151.

[18] M. D. Gunzburger, L. S. Hou, and T. Svobodny, *Analysis and finite element approximation of optimal control problems for the stationary Navier–Stokes equations with Cirichlet controls*, RAIRO Model. Math. Anal. Numer., **25** (1991), 711–748.

[19] M. D. Gunzburger and L. S. Hou, *Finite-dimensional approximation of a class of constrained nonlinear optimal control problems*, SIAM J. Control. Optim., **34** (1996), 1001–1043.

[20] L. S. Hou and S. S. Lavindran, *A penalized neumann control approach for solving an optimal dirichlet control problem for the Navier–Stokes equations*, SIAM J. Control. Optim., **36** (1998), 1795–1814.

[21] I. Babuska, B. A. Szabo, and I. N. Katz, *The p-version of the finite element method*, SIAM J. Numer. Anal., **18** (1981), 515–545.

[22] W. Gui and I. Babuska, *The h, p and h-p versions of the finite element method in 1 dimension. I. The error analysis of the p-version*, Numer. Math., **49** (1986), 577–612.

[23] W. Gui and I. Babuska, *The h, p and h-p versions of the finite element method in 1 dimension. II. The error analysis of the h- and h-p versions*, Numer. Math., **49** (1986), 613–657.

[24] W. Gui and I. Babuska, *The h, p and h-p versions of the finite element method in 1 dimension. III. The adaptive h-p version*, Numer. Math., **49** (1986), 659–683.

[25] B. Guo and I. Babuska, it The h-p version of the finite element method. I. The basic approximation results, Comput. Mech., **1** (1986), 21–41.

[26] I. Babuska and B. Q. Guo, *The h-p version of the finite element method for domains with curved boundaries*, SIAM J. Numer. Anal., **25** (1988), 837–861.

[27] B. Guo, *The h-p version of the finite element method for elliptic equations of order 2m*, Numer. Math., **53** (1988), 199–224.

[28] I. Babuska and H.-S. Oh, *The p-version of the finite element method for domains with corners and for infinite domains*, Numer. Methods Partial Differential Eq., **6** (1990), 371–392.

[29] I. Babuska and M. Suri, *The p and h-p versions of the finite element method*, basic principles and properties, SIAM Rev., **36** (1994), 578–632.

[30] I. Babuska and J. M. Melenk, *The partition of unity finite element method*, Technical Note BN-1185, Inst. Phys. Sci. Tech., (1995).

[31] I. Babuska and F. Ihlenburg, *Finite element solution of the Helmholtz equation with high wave number part II: The h-p version of the FEM*, SIAM J. Numer. Anal., **34** (1997), 315–358.

[32] P. Chen and A. Quarteroni, *Weighted Reduced Basis Method for Stochastic Optimal Control Problems with Elliptic PDE Constraint*, SIAM/ASA J. Uncert. Quant., **2** (2014), 364–396.

[33] E. Rosseel and G. N. Wells, *Optimal control with stochastic PDE constraints and uncertain controls*, Comput. Methods Appl. Mech. Engrg., (2012), 152–167.

[34] H. Tiesler, R. M. Kirby, D. Xiu, and T. Preusser, *Stochastic collocation for optimal control problems with stochastic PDE constraints*, SIAM J. Control Optim., **50** (2012), 2659–2682.

[35] M. Eigel, C. J. Gittelson, C. Schwab, and E. Zander, *A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes*, ESAIM: M2AN, **49** (2015), 1367–1398.

[36] R. Adams, *Sobolev Spaces*, Academic, New York, 1975.

[37] J. Galvis and M. Sarkis, *Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity*, SIAM J. Numer. Anal., **47** (2009), 3624–3651.

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

[39] R. G. Ghanem and P. D. Spanos, *Stochastic finite elements: A spectral approach*, Springer-Verlag, 1991.

[40] W. Luo, *Wiener chaos expansion and numerical solutions of stochastic partial differential equations*, Ph.D. thesis, California institute of technology, Pasadena, California 2006.

[41] F. Brezzi, J. Rappaz, and P. Raviart, *Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions*, Numer. Math., **36** (1980), 1–25.

[42] M. Crouzeix and J. Rappaz, *On numerical approximation in bifurcation theory*, Masson, Parix, 1990.

[43] V. Girault and P. Raviart, *Finite element methods for Navier–Stokes equations*, Springer, Berlin, 1986.

[44] D. Xiu, *Numerical methods for stochastic computations*s, Princeton, 2010.