Commun. Comput. Phys.
doi: 10.4208/cicp.241011.150612a

Vol. **14**, No. 1, pp. 77-106
July 2013

A Stochastic Galerkin Method for Stochastic Control Problems

Hyung-Chun Lee[1] and Jangwoon Lee[2,*]

[1] *Department of Mathematics, Ajou University, Suwon, Korea 443-749.*
[2] *Department of Mathematics, University of Mary Washington, Fredericksburg, VA 22401, USA.*

Received 24 October 2011; Accepted (in revised version) 15 June 2012

Available online 18 October 2012

Abstract. In an interdisciplinary field on mathematics and physics, we examine a physical problem, fluid flow in porous media, which is represented by a stochastic partial differential equation (SPDE). We first give a priori error estimates for the solutions to an optimization problem constrained by the physical model under lower regularity assumptions than the literature. We then use the concept of Galerkin finite element methods to establish a new numerical algorithm to give approximations for our stochastic optimal physical problem. Finally, we develop original computer programs based on the algorithm and use several numerical examples of various situations to see how well our solver works by comparing its outputs to the priori error estimates.

AMS subject classifications: 65M55, 65N30

Key words: Stochastic Galerkin methods, stochastic partial differential equation, distributed control, finite element methods.

1 Introduction

In the last decade, people in the scientific computing community have taken great interest in the stochastic partial differential equations (SPDEs) and its solver called the Stochastic Galerkin Method [4, 5, 21, 29, 33]. In this paper, we use the idea from the Galerkin finite element method to analyze optimal control problems constrained by SPDE and develop its numerical solver.

The stochastic Galerkin method has been created and developed to analyze a stochastic problem in the following sense. Suppose that we have a deterministic partial differential equation (PDE) that models some natural phenomenon; for instance, pollutant transportation in groundwater. To improve this deterministic mathematical model, we assume

*Corresponding author. *Email addresses:* `hclee@ajou.ac.kr` (H.-C. Lee), `llee3@umw.edu` (J. Lee)

that we replace some deterministic quantities in the PDE with stochastic input data. For example, there may be lack of knowledge about some materials such as rocks and soils for groundwater. For these unknown properties of rocks and/or soils, we would like to use the concept of randomness in the model so that a new mathematical model with additional random terms can represent better the natural phenomenon. If there are inputs that are random, then the solution to the new model problem should also be including randomness. We then need a stochastic domain, and may need to use probability theories to analyze the solution to the new model problem. Now the remaining question is how we apply or modify a typical method such as the Galerkin Method to analyze the new stochastic problem derived from a deterministic problem. The stochastic Galerkin method actually answers this question, and it turns out to be a good method that requires less computational efforts than Monte Carlo method in computing $E[u]$ for sufficiently strict accuracy requirements (see [4]). However, in case that one use many terms in the K-L expansion of our coefficient a, the Monte Carlo method is known to be most effective (see [5]).

In this paper, we analyze the stochastic optimal control problems subject to SPDE by using a similar approach in the literature that is used to solve SPDEs. For example, we use the truncated Karhunen-Loéve (K-L) expansion as a main tool to convert the stochastic optimal control problem to a coupled optimality system of deterministic PDEs. In fact, in the last decade, based on the K-L expansion, there has been much progress in both the analysis and the finite element approximations of SPDEs; see, e.g., [2–5, 14, 21, 29, 33, 40].

Notwithstanding the many papers devoted to discrete approximations of solutions of SPDEs and optimal control problems for SPDEs, the literature seems to lack $L^2(\Gamma;H_0^1(D))$ convergence results for optimal distributed control problems based on the K-L expansion with both feasibility and efficiency of rigorous error analysis demonstrated via numerical examples. The goals of this work are to establish the convergence of the solution of a distributed optimal control problem with the K-L expansions, derive its error estimates in the norm of the solution space under minimal regularity assumptions in the y-direction, and show the practicability and effectiveness of our theories using numerical examples.

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:

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

where $\mathbb{E}$ denotes expected value, D the spatial domain, ∂D its boundary, U a target solution to the constraint, β a positive constant that says the importance between two terms in (1.1), and f a deterministic control acting in the domain. Here, our stochastic elliptic PDE generally models fluid flow in porous media. Under the homogeneous Dirichlet boundary condition, for almost every $\omega \in \Omega$, we look for a solution u, stochastic function from $\overline{D} \times \Omega$ to $\mathbb{R}$: where $D \subset \mathbb{R}^d$ is a convex bounded polygonal domain, $a: D \times \Omega \to \mathbb{R}$ is a stochastic function with a bounded, continuous covariance function (this is for the KL expansion) and a uniformly bounded, continuous first derivative (this is for the regularity of the solution u), and $f \in L^2(D)$ is a distributed deterministic control. Note that in the paper, ∇ means differentiation with respect to $x \in D$ only.

To analyze this stochastic optimal control problem, we first estimate the error of the solution to SPDE, and then use the Brezzi-Rappaz-Raviart (BRR) theory (this idea was first applied to analyze constrained nonlinear optimal control problems in [28], and then used many times in optimal control problems subject to PDEs; see. e.g., [11, 26, 27, 31]) in uncoupling the optimality system of equations that eventually gives us the error estimate of the solution to the stochastic optimal control problem. Then at the end, we construct a computational algorithm for our stochastic control problem and present some numerical examples with a given target solution to the stochastic optimal control problem with a distributed control in the domain.

Some remarks about the literature are in order. Originally, Ghanem and Spanos proposed the stochastic Galerkin method as a spectral discretization technique in the random dimension ([23]). Recently, Stefanou [41] provided a review paper that summarized past and recent developments as well as future directions in the stochastic finite element method. In [5, 7], as recent trends in developing the stochastic Galerkin method, the authors of both papers converted the stochastic problem to a sequence of N-dimensional (here, N is the number of terms in the truncated K-L expansion), parametric deterministic problems. The paper [17] applied an approximate K-L expansion to the solution to reduce the set of deterministic PDEs, addressing computational efficiency of the spectral stochastic Galerkin method to SPDEs.

The work [29] shows that the authors analyzed the stochastic optimal control problem subject to the elliptic SPDE under the Neumann boundary condition; the control used was of the deterministic, boundary-value type. In [33], finite element methods for a standard stochastic elliptic PDE-constrained optimal control problem was considered; the Wiener-Ito (W-I) chaos expansion was used for the diffusion coefficient as a main analytical tool; this shows some results in continuously differentiable solution spaces which require more regularity assumptions. Unlike the references [2–5, 14, 21, 40] related to this paper, in [30], a different notion of SPDEs was provided using the W-I chaos expansion and the Wick product based on a different physical modeling situation, e.g., the expectation of the Wick product of a and u is the Wick product of the expectations of a and u, which is not true in general for the standard product. In [37], using the Wick product properties, a finite element approximation of the linear SPDEs driven by a multiplicative white noise was presented. A key idea used there is to reformulate the SPDEs as an infinite set of deterministic PDEs with the Wick product. Also, in [38], optimal control problems for linear SPDEs with quadratic cost functionals and distributed stochastic controls were considered using the Galerkin finite element method; the W-I chaos expansions for both the solution and control were used. Recently, the work [39] proposed a stochastic finite element method to the optimal control problem. It also provided a number of numerical examples to demonstrate the proposed formulation.

In [10], Loeb-space methods were used to prove the existence of an optimal control for the two-dimensional stochastic Navier-Stokes equations in a variety of settings including that of control based on digital observations of the evolution of the solution. In [16], the authors consider a control problem for a stochastic Burgers equation. They studied a sequence of approximated Hamilton-Jacobi equations by using dynamic programming. In [15], the purpose was not only to prove existence of optimal controls but mainly to characterize them by an optimal feedback law, i.e., they wished to perform the standard program of synthesis of the optimal control that consists in the following steps: first they solved (in a suitable sense) the Hamilton-Jacobi-Bellman equation; then they proved that such a solution was the value function of the control problem and allowed them to construct the optimal feedback law.

The plan of the paper is as follows. In Section 2, we represent a random field, introducing the K-L expansion and its truncated expansion. In Section 3, we analyze our constraint equation, stochastic elliptic PDE, transforming a stochastic problem to a high dimensional deterministic problem and presenting a priori error estimates for the equation. In Section 4, we derived the optimality system of equations, showing the existence of a Lagrange multiplier. Then in Section 5, we establish error estimate for the discrete approximations of solutions to the optimality system. Finally in Section 6, we construct a mathematical algorithm and give several numerical examples of stochastic optimal control problems constrained by the stochastic elliptic PDE under the Dirichlet boundary condition.

2 Preliminaries and function spaces

2.1 Karhunen-Loève expansions

In this section, we introduce the K-L expansions, which is well known as a theoretical tool for approximating stochastic functions; see [5,21,23,36]. If $a(x,\omega)$ is a stochastic function that has a continuous and bounded covariance function, it can be represented by

$$a(x,\omega)=\mathbb{E}a(x,\omega)+\sum_{n\geq 1}\sqrt{\lambda_n}\phi_n(x)X_n(\omega), \tag{2.1}$$

where $\mathbb{E}X_n(\omega)=0$, $\mathbb{E}(X_n(\omega)X_m(\omega))=\delta_{nm}$, and $(\lambda_n,\phi_n(x))$ are solutions to the eigenvalue problem:

$$\int_D C(x_1,x_2)\phi_n(x_1)\,dx_1=\lambda_n\phi_n(x_2), \tag{2.2}$$

where $C(x_1,x_2)=\mathbb{E}(a(x_1,\omega)a(x_2,\omega))-\mathbb{E}a(x_1,\omega)\mathbb{E}a(x_2,\omega)$. We call this expansion the K-L expansion of $a(x,\omega)$.

Whenever we use numerical methods to approximate the solutions of mathematical models, we use only finitely terms in expansions. Also, in realistic mathematical models, it is known that the random source can be expressed by finitely many random variables. For those reasons, for the use of KL expansions in numerical methods, we define truncated KL expansions from (2.1):

$$a_N(x,\omega) = \mathbb{E}a(x,\omega) + \sum_{n=1}^{N} \sqrt{\lambda_n}\phi_n(x)X_n(\omega). \tag{2.3}$$

Remark 2.1. Later, for the existence of the solution to our stochastic PDE, we shall require a to be between two positive constants; i.e., $m \le a(x,\omega) \le M$. For this condition, as a practical example, a could have a log normal distribution (see [22]).

Not to have modeling errors in computation, it is natural to ask the convergence of the truncated KL expansions. As such, we state the convergence theorem based on Mercer's theorem for the KL expansions.

Proposition 2.1. The truncated KL expansion $a_N(x,\omega)$of a stochastic function $a(x,\omega)$ converges uniformly to $a(x,\omega)$

$$\sup_{x\in D}\mathbb{E}[(a(x,\omega)-a_N(x,\omega))^2] = \sup_{x\in D}\sum_{n=N+1}^{\infty}\lambda_n\phi_n^2(x) \to 0 \quad \text{as } N\to\infty. \tag{2.4}$$

The detailed convergence results can be found in the literature (e.g., see [4]).

2.2 Function spaces and notation

For our stochastic elliptic problems, we use a complete probability space $(\Omega,\mathcal{F},P)$, 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.

We use standard Sobolev space notation (see [1]). For instance, $H^1(D)$ is a Hilbert space with a norm $\|\cdot\|_{H^1(D)}$; $H_0^1(D)$ is the subspace of $H^1(D)$ whose function value is zero on the boundary of D, and its norm is $\|u\|^2_{H_0^1(D)} = \int_D |\nabla u|^2\, dx$.

With these standard Sobolev spaces, we define stochastic Sobolev spaces as follows:

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

where

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

Similarly, we can define $L^2(\Omega;L^2(D))$. For simplicity, we let $\mathcal{L}^2(D) = L^2(\Omega;L^2(D))$ and $\mathcal{H}_0^1(D) = L^2(\Omega;H_0^1(D))$. Note that these stochastic Sobolev spaces are Hilbert spaces.

For the weak formulation of our stochastic elliptic PDE, we introduce the following notations:

$$b[u,v] = \mathbb{E}\int_D a\nabla u\cdot\nabla v\, dx \tag{2.5}$$

and

$$[u,v] = \mathbb{E}\int_D uv\, dx, \tag{2.6}$$

where $\mathbb{E}$ is the expected value.

2.3 Existence and uniqueness of the solution

Using notations (2.5) and (2.6) introduced in the previous section, we can derive the weak formulation of the strong formulation (1.2): seek $u \in \mathcal{H}_0^1(D)$ such that

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

In this paper, to have the existence and uniqueness of the solution to our stochastic elliptic problems (1.2), we assume that there are positive m and M such that

$$m \leq a(x,\omega) \leq M \quad \text{a.e. } (x,\omega) \in D \times \Omega. \tag{2.7}$$

Remark 2.2. For the condition (2.7) for $a(x,\omega)$, as a practical example, a could have a log normal distribution (see [22]).

Then from the Lax-Milgram lemma (see [8]), we have the following theorem about the existence and uniqueness of the solution:

Theorem 2.1. *Let $f \in L^2(D)$. Then there is a unique solution to the following weak formulation: find $u \in \mathcal{H}_0^1(D)$ such that*

$$b[u,v] = [f,v] \quad \forall v \in \mathcal{H}_0^1(D). \tag{2.8}$$

Proof. Note that from ellipticity condition (2.7), there exists $M,m > 0$ such that

$$|b[u,v]| \leq M \|u\|_{\mathcal{H}_0^1(D)} \|v\|_{\mathcal{H}_0^1(D)} \quad \forall u,v \in \mathcal{H}_0^1(D)$$

and

$$m\|v\|^2_{\mathcal{H}_0^1(D)} \leq b[v,v] \quad \forall v \in \mathcal{H}_0^1(D).$$

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

$$|[f,v]| \leq C\|v\|_{\mathcal{H}_0^1(D)}$$

for any $v \in \mathcal{H}_0^1(D)$. Hence, by the Lax-Milgram lemma (cf. [8]), (2.8) has a unique solution. $\square$

Then from standard arguments in measure theory, we can show that the solution to the weak formulation (2.8) solves our stochastic PDE (1.2) under the Dirichlet condition (e.g., see [25]).

3 Models with finite dimensional information

As we discussed before, both because the random sources in realistic models can be expressed by finitely many mutually independent random variables and because infinite expansions should be handled by finite expansions in numerical methods, we assume that we have finite dimensional information on $a(x,\omega)$:

$$a(x,\omega)=\mathbb{E}a(x,\omega)+\sum_{n=1}^{N}\sqrt{\lambda_n}\phi_n(x)X_n(\omega); \tag{3.1}$$

i.e., from now on, our coefficient $a(x,\omega)$ is the same as the truncated one $a_N(x,\omega)$ (see (2.3)).

Remark 3.1. For the truncated assumption (3.1), as was discussed in [4], assumption (3.1) may be valid in its own right in practical applications; also, similar to the convergence analysis of [4] for the truncated problem, we may prove the convergence of the truncated control problem based on Mercer's theorem.

Because our assumption (2.7) on $a(x,\omega)$ does not automatically imply the boundedness of the truncated KL expansion (3.1), to have the existence and uniqueness of the solution for our models with finite dimensional information, it is necessary that $\mathbb{E}a(x,\omega)+\sum_{n=1}^{N}\sqrt{\lambda_n}\phi_n(x)X_n(\omega)$ satisfy a similar condition (2.7); i.e., we assume that there exist $m,M>0$ such that

$$m\le\mathbb{E}a(x,\omega)+\sum_{n=1}^{N}\sqrt{\lambda_n}\phi_n(x)X_n(\omega)\le M \quad \text{a.e. } (x,\omega)\in D\times\Omega. \tag{3.2}$$

Remark 3.2. For the ellipticity assumption (3.2), we refer the readers to [22] for a discussion of the ellipticity condition for $a(x,\omega)$, where $\log a(x,\omega)$ is a Gaussian field.

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 $y\in\Gamma\equiv\prod_{n=1}^{N}\Gamma_n\subset\mathbb{R}^N$ of $(X_1,X_2,\cdots,X_N)$. Under these assumptions, the solution of (2.8) 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., [2,4,14].

Under the above assumptions, we have the following high-dimensional deterministic equivalent weak formulation of (2.8) with the finite dimensional information:

$$\int_\Gamma\rho(y)\int_D a(x,y)\nabla u(x,y)\cdot\nabla v(x,y)\;dxdy=\int_\Gamma\rho(y)\int_D f(x)v(x,y)\;dxdy. \tag{3.3}$$

The strong formulation of (3.3) is

$$-\nabla\cdot[a(x,y)\nabla u(x,y)]=f(x),\quad \forall(x,y)\in D\times\Gamma, \tag{3.4a}$$

$$u(x,y)=0,\qquad \forall(x,y)\in\partial D\times\Gamma. \tag{3.4b}$$

Then we have well-posedness of (3.4) because a is bounded. Also, by using the finite element method, we can provide the solution to a SPDE from solving (3.4).

For the high-dimensional elliptic PDE, we recall Sobolev spaces as follows:

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

where

$$\|v\|^2_{L^2(\Gamma;H_0^1(D))}=\int_\Gamma\rho\,\|v\|^2_{H_0^1(D)}\,dy=\mathbb{E}\|v\|^2_{H_0^1(D)}.$$

Similarly, we can define $L^2(\Gamma;L^2(D))$. For simplicity, we set $\mathcal{L}^2(D)=L^2(\Gamma;L^2(D))$ and $\mathcal{H}_0^1(D)=L^2(\Gamma;H_0^1(D))$ likewise before.

Corresponding notations are

$$b[u,v]=\int_\Gamma\rho\int_D a\nabla u\cdot\nabla v\,dxdy\qquad\text{and}\qquad[u,v]=\int_\Gamma\rho\int_D uv\,dxdy.$$

3.1 Finite element spaces

Let us first consider finite element spaces on $D\subset\mathbb{R}^d$. Let X^h and G^h be families of finite element approximation subspaces of $H_0^1(D)$ and $L^2(D)$ that consist of piecewise linear continuous functions defined over a family of regular triangulations of D with a maximum grid size parameter $h>0$. We assume that X^h and G^h satisfy the following approximation properties:

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

$$\inf_{\phi^h\in X^h}\|\phi-\phi^h\|_{H_0^1(D)}\le Ch\|\phi\|_{H^2(D)};\tag{3.5}$$

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

$$\inf_{\phi^h\in G^h}\|\phi-\phi^h\|_{L^2(D)}\le Ch\|\phi\|_{H_0^1(D)}.\tag{3.6}$$

Next, we consider finite element spaces on $\Gamma\subset\mathbb{R}^N$. We partition Γ into a finite number of disjoint $\mathbb{R}^N$ boxes B_i^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 $B_k^N\cap B_l^N=\emptyset$ for $k\neq l\in I$ and $(a_i^j,\,b_i^j)\subset\Gamma_j$.

A maximum grid size parameter $\delta>0$ is denoted by

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

Let $Y^\delta\subset L^2(\Gamma)$ be the finite element approximation space of piecewise polynomials with degree at most p_j on each direction y_j. Thus if $\psi^\delta\in Y^\delta$, then $\psi^\delta|_{B_i^N}\in$ span$(\prod_{j=1}^N y_j^{n_j}:$

$n_j \in \mathbb{N}$ and $n_j \le p_j$). Letting $p = (p_1, p_2, \cdots, p_N)$, we have (cf. see [8]) the following property: for all $\psi \in H^{p+1}(\Gamma)$,

$$\inf_{\psi^\delta \in Y^\delta} \|\psi - \psi^\delta\|_{L^2(\Gamma)} \le \delta^\gamma \sum_{j=1}^{N} \frac{\|\partial_{y_j}^{p_j+1} \psi\|_{L^2(\Gamma)}}{(p_j+1)!}, \tag{3.7}$$

where $\gamma = \min_{1\le j\le N}\{p_j+1\}$ if $\delta < 1$ and $\gamma = \max_{1\le j\le N}\{p_j+1\}$ otherwise.

We now think of finite element spaces on $D\times\Gamma$, say $V^{h\delta}$. Here, if $v^{h\delta} \in V^{h\delta}$, $v^{h\delta} \in$ span($\phi^h\psi^\delta$: $\phi^h(x) \in X^h$ and $\psi^\delta(y) \in Y^\delta$).

We denote by R^h the $H^1(D)$-projection from $H_0^1(D)$ onto X^h and P^δ the $L^2(\Gamma)$-projection from $L^2(\Gamma)$ onto Y^δ. Namely for each $\phi \in H_0^1(D)$,

$$(R^h\phi, \phi^h)_{H_0^1(D)} = (\phi, \phi^h)_{H_0^1(D)} \quad \forall \phi^h \in X^h;$$

for each $\psi \in L^2(\Gamma)$,

$$(P^\delta\psi, \psi^\delta)_{L^2(\Gamma)} = (\psi, \psi^\delta)_{L^2(\Gamma)} \quad \forall \psi^\delta \in Y^\delta.$$

It follows from (3.5) that for all $\phi \in H_0^1(D) \cap H^2$ and for some $C > 0$, we have

$$\|\phi - R^h\phi\|_{H_0^1(D)} \le Ch\|\phi\|_{H^2(D)} \tag{3.8}$$

and from (3.7), for all $\psi \in H^{p+1}(\Gamma)$, we obtain

$$\|\psi - P^\delta\psi\|_{L^2(\Gamma)} \le \delta^\gamma \sum_{j=1}^{N} \frac{\|\partial_{y_j}^{p_j+1}\psi\|_{L^2(\Gamma)}}{(p_j+1)!}. \tag{3.9}$$

By using the last two inequalities, we have the following property; see [4]: for all $u \in C^{p+1}(\Gamma; H^2(D) \cap H_0^1(D))$, there exists $C > 0$, which is independent of h, δ, N, and p, such that

$$\inf_{u^{h\delta} \in V^{h\delta}} \|u - u^{h\delta}\|_{\mathcal{H}_0^1(D)} \le C\left(h\|u\|_{\mathcal{H}^2(D)} + \delta^\gamma \sum_{j=1}^{N} \frac{\|\partial_{y_j}^{p_j+1} u\|_{\mathcal{H}_0^1(D)}}{(p_j+1)!} \right). \tag{3.10}$$

Remark 3.3. Because $a(x,y) = \mathbb{E}a(x,y) + \sum_{n=1}^{N} \sqrt{\lambda_n}\phi_n(x) y_n \in C^{p+1}(\overline{D\times\Gamma})$, it is well known that the solution u of (3.4) satisfies $u \in C^{p+1}(\Gamma; H^2(D)\cap H_0^1(D))$; see e.g., Lemma 4.1 in [34] and Remark 5.1 in [5]. Also if we assume that $f(x,y) \in C^{p+1}(\Gamma; L^2(D))$ (i.e., if f has a finite K-L expansion) to carry out the analysis of the finite element method, then the solution u of the following problem also satisfies $u \in C^{p+1}(\Gamma; H^2(D)\cap H_0^1(D))$:

$$-\mathrm{div}[a(x,y)\nabla u(x,y)] = f(x,y), \quad \forall (x,y) \in D\times\Gamma, \tag{3.11a}$$

$$u(x,y) = 0, \quad \forall (x,y) \in \partial D\times\Gamma. \tag{3.11b}$$

Hence, under our assumptions, (3.10) makes sense.

3.2 Error estimates for high dimensional elliptic PDEs

Recall that our goal is to solve (3.4). The stochastic weak formulation of (3.4) is as follows: seek $u\in\mathcal{H}_0^1(D)$ such that for all $v\in\mathcal{H}_0^1(D)$,

$$b[u,v]=[f,v]. \tag{3.12}$$

Then we have the finite element weak formulation: find $u^{h\delta}\in V^{h\delta}$ such that for all $v^{h\delta}\in V^{h\delta}$,

$$b[u^{h\delta},v^{h\delta}]=[f,v^{h\delta}]. \tag{3.13}$$

We want error estimate of solutions for (3.12) and (3.13) in $\mathcal{H}_0^1(D)$. Also we do the same thing with a finite data $g(x,y)$ instead of $f(x)$. For these, we consider the following lemmas.

Lemma 3.1. *Let $f(x)\in L^2(D)$. Then for any $y\in\Gamma$, $u(\cdot,y)\in H^2(D)$ and there exists $C>0$ such that*

$$\|u(\cdot,y)\|_{H^2(D)}\le C\|f\|_{L^2(D)}.$$

Proof. The proof of this lemma can be found in [18], for instance. □

Remark 3.4. For problems with $g(\cdot,y)\in L^2(D)$, we have

$$\|u(\cdot,y)\|_{H^2(D)}\le C\|g(\cdot,y)\|_{L^2(D)}.$$

Lemma 3.2. *Let $f(x)\in L^2(D)$ and $\phi_j(x)\in L^\infty(D)$. Then for all $j=1,2,\cdots,N$ and for any $y\in\Gamma$, there exists $C>0$ such that*

$$\frac{\|\partial_{y_j}^{p_j+1}u(\cdot,y)\|_{H_0^1(D)}}{(p_j+1)!}\le C\|\phi_j\|_{L^\infty(D)}^{p_j+1}\|f\|_{L^2(D)}.$$

Proof. Without loss of generality, we show this for only $j=1$. Recall that $a(x,y)=\mathbb{E}a(x,y)+\sum_{j=1}^N\sqrt{\lambda_j}\phi_j(x)y_j=\sum_{j=0}^N\sqrt{\lambda_j}\phi_j(x)y_j$ with $\lambda_0=1=y_0$ and $\phi_0(x)=\mathbb{E}a(x,y)$. For convenience, we set

$$a(x,y)=\sum_{j=0}^N\phi_j(x)y_j$$

with redefined $\phi_j(x)=\sqrt{\lambda_j}\phi_j(x)$.

If we take derivatives with respect to y_1 in (3.4), we find

$$-\nabla\cdot[\phi_1(x)\nabla u(x,y)+a(x,y)\nabla\partial_{y_1}u(x,y)]=0.$$

Note that because $u(x,y)=0$ for any $(x,y)\in\partial D\times\Gamma$, $\partial_{y_1}u(x,y)=0$ for any $(x,y)\in\partial D\times\Gamma$. Thus, by integrating over D after multiplying by $\partial_{y_1}u$, we see that

$$\int_D\phi_1(x)\nabla u(x,y)\cdot\nabla\partial_{y_1}u(x,y)\,dx+\int_D a(x,y)|\nabla\partial_{y_1}u(x,y)|^2\,dx=0.$$

This by the coercivity, implies

$$\|\partial_{y_1} u(\cdot,y)\|^2_{H^1_0(D)} \le C\|\phi_1\|_{L^\infty(D)}\|u(\cdot,y)\|_{H^1(D)}\|\partial_{y_1} u(\cdot,y)\|_{H^1(D)}.$$

From Lemma 3.1, we have

$$\|\partial_{y_1} u(\cdot,y)\|_{H^1_0(D)} \le C\|\phi_1\|_{L^\infty(D)}\|f\|_{L^2(D)}.$$

We now assume that the following is true:

$$\|\partial^{p_1}_{y_1} u(\cdot,y)\|_{H^1_0(D)} \le Cp_1!\|\phi_1\|^{p_1}_{L^\infty(D)}\|f\|_{L^2(D)}. \tag{3.14}$$

Taking derivatives p_1+1 times with respect to y_1 in (3.4), we obtain

$$-\nabla\cdot[(p_1+1)\phi_1(x)\nabla\partial^{p_1}_{y_1}u(x,y)+a(x,y)\nabla\partial^{p_1+1}_{y_1}u(x,y)]=0.$$

After multiplying by $\partial^{p_1+1}_{y_1}u$ integrating over D yields

$$(p_1+1)\int_D \phi_1(x)\nabla\partial^{p_1}_{y_1}u(\cdot,y)\cdot\nabla\partial^{p_1+1}_{y_1}u(\cdot,y)\,dx+\int_D a(x,y)|\nabla\partial^{p_1+1}_{y_1}u(\cdot,y)|^2\,dx=0.$$

By the coercivity, Lemma 3.1, and our induction hypothesis (3.14), we find

$$\|\partial^{p_1+1}_{y_1}u(\cdot,y)\|_{H^1_0(D)} \le C(p_1+1)\|\phi_1\|_{L^\infty(D)}(p_1!\|\phi_1\|^{p_1}_{L^\infty(D)}\|f\|_{L^2(D)}).$$

Thus, the assertion for $j=1$ follows from the last inequality by induction. □

Remark 3.5. For problems with $g(x,y)\in C^{p+1}(\Gamma;L^2(D))$, we have

$$\frac{\|\partial^{p_j+1}_{y_j}u(\cdot,y)\|_{H^1_0(D)}}{(p_j+1)!} \le C\sum_{k=0}^{p_j+1}\|\phi_j\|^{p_j+1-k}_{L^\infty(D)}\|\partial^k_{y_j}g(\cdot,y)\|_{L^2(D)}.$$

We now have the following theorem.

Theorem 3.1. *Let $f(x)\in L^2(D)$, u be the solution of* (3.12)*, and $u^{h\delta}$ be the finite element solution of* (3.13)*. Then there exists $C>0$ such that*

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

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

Proof. As a consequence of (3.10), Lemma 3.1, and Lemma 3.2, we have

$$\begin{aligned}\|u-u^{h\delta}\|_{\mathcal{H}_0^1(D)} &\le C\left(h\|u\|_{\mathcal{H}^2(D)}+\delta^\gamma\sum_{j=1}^N\frac{\|\partial_{y_j}^{p_j+1}u\|_{\mathcal{H}_0^1(D)}}{(p_j+1)!}\right)\\ &\le C\left(h\|f\|_{L^2(D)}+\delta^\gamma\sum_{j=1}^N\|\phi_j\|_{L^\infty(D)}^{p_j+1}\|f\|_{L^2(D)}\right)\\ &\le C(h+\delta^\gamma)K\|f\|_{L^2(D)},\end{aligned}$$

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

Similarly, (3.10), Remark 3.4, and Remark 3.5 give the following remark.

Remark 3.6. For problems with $g(x,y)\in C^{p+1}(\Gamma;L^2(D))$, we have

$$\|u-u^{h\delta}\|_{\mathcal{H}_0^1(D)}\le C(h+\delta^\gamma)K\|g\|_{\mathcal{L}^2(D)}, \tag{3.15}$$

where

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

4 Stochastic distributed control problems

4.1 Existence of an optimal solution

In this section, we examine the existence of an optimal solution that minimizes our functional (1.1). We first define

$$\mathcal{U}_{ad}=\{(u,f)\in\mathcal{H}_0^1\times L^2 \text{ such that (2.8) satisfied and } \mathcal{J}(u,f)<\infty\} \tag{4.1}$$

be the admissibility set. Then $(\hat{u},\hat{f})\in\mathcal{U}_{ad}$ is said to be an *optimal solution* of $\mathcal{J}(u,f)$ if for all $(u,f)\in\mathcal{U}_{ad}$ satisfying that $\|u-\hat{u}\|_{\mathcal{H}_0^1(D)}+\|f-\hat{f}\|_{L^2(D)}\le\epsilon$ for some $\epsilon>0$,

$$\mathcal{J}(\hat{u},\hat{f})\le\mathcal{J}(u,f). \tag{4.2}$$

We now prove the existence of an optimal solution.

Theorem 4.1. *There is an optimal solution $(\hat{u},\hat{f})\in\mathcal{U}_{ad}$ of $\mathcal{J}(u,f)$.*

Proof. We know that $\mathcal{U}_{ad}$ is not empty. We consider a minimizing sequence $\{(u^{(n)},f^{(n)})\}$ in $\mathcal{U}_{ad}$ such that

$$\lim_{n\to\infty}\mathcal{J}(u^{(n)},f^{(n)})=\inf_{(u,f)\in\mathcal{U}_{ad}}\mathcal{J}(u,f). \tag{4.3}$$

Then the sequence $\{f^{(n)}\}$ is uniformly bounded in $L^2(D)$. This implies that the sequence $\{u^{(n)}\}$ is also uniformly bounded in $\mathcal{H}_0^1(D)$. As a result, there is a subsequence $\{(u^{(n_i)},f^{(n_i)})\}$ of $\{(u^{(n)},f^{(n)})\}$ that weekly converge. That is, there exists $(\hat{u},\hat{f})\in\mathcal{H}_0^1(D)\times L^2(D)$ such that

$$u^{(n_i)}\rightharpoonup\hat{u}\ \text{ weakly in } \mathcal{H}_0^1(D) \quad \text{and} \quad f^{(n_i)}\rightharpoonup\hat{f}\ \text{ weakly in } L^2(D). \tag{4.4}$$

Note that $\{f^{(n)}\}$ is uniformly bounded in $\mathcal{L}^2(D)$. Thus, we have

$$f^{(n_i)}\rightharpoonup\hat{f}\ \text{ weakly in } \mathcal{L}^2(D). \tag{4.5}$$

This implies that

$$[f^{(n_i)},v]\rightarrow[\hat{f},v] \quad \forall v\in\mathcal{L}^2(D). \tag{4.6}$$

Also note that $\{\nabla u^{(n)}\}$ is also uniformly bounded in $\mathcal{L}^2(D)$. Thus, we have

$$\nabla u^{(n_i)}\rightharpoonup\nabla\hat{u} \quad \text{weakly in } \mathcal{L}^2(D).$$

This yields

$$[\nabla u^{(n_i)},w]\rightarrow[\nabla\hat{u},w] \quad \forall w\in\mathcal{L}^2(D).$$

The fact that $\nabla v\in\mathcal{L}^2(D)$ for $v\in\mathcal{H}_0^1(D)$ lead us to

$$[\nabla u^{(n_i)},\nabla v]\rightarrow[\nabla\hat{u},\nabla v] \quad \forall v\in\mathcal{H}_0^1(D).$$

Because $a\nabla v\in\mathcal{L}^2(D)$ for $v\in\mathcal{H}_0^1(D)$, we obtain

$$b[u^{(n_i)},v]\rightarrow b[\hat{u},v] \quad \forall v\in\mathcal{H}_0^1(D). \tag{4.7}$$

With the help of (4.6) and (4.7), we can show that

$$b[\hat{u},v]=\lim_{n_i\to\infty}b[u^{(n_i)},v]=\lim_{n_i\to\infty}[f^{(n_i)},v]=[\hat{f},v] \quad \forall v\in\mathcal{H}_0^1(D). \tag{4.8}$$

That is, $(\hat{u},\hat{f})$ satisfies (2.8) and hence $(\hat{u},\hat{f})\in\mathcal{U}_{ad}$. Using the weak convergence (4.4) and the weak lower continuity of the functional $\mathcal{J}(\cdot,\cdot)$ we arrive at

$$\mathcal{J}(\hat{u},\hat{f})\leq\lim_{n_i\to\infty}\inf\mathcal{J}(u^{(n_i)},f^{(n_i)})=\inf_{(u,f)\in\mathcal{U}_{ad}}\mathcal{J}(u,f). \tag{4.9}$$

Therefore, $(\hat{u},\hat{f})$ is an optimal solution. $\square$

4.2 The optimality system of stochastic equations

We will derive an optimality system of stochastic equations by using the Lagrange multiplier rule for the constrained minimization problem:

$$\min_{(u,f)\in\mathcal{U}_{ad}} \mathcal{J}(u,f) \quad \text{subject to} \quad (2.8). \tag{4.10}$$

For the deterministic case, we know that there exists a Lagrange multiplier; see e.g., [18]. Thus, without proving the existence of a Lagrange multiplier, we could derive an optimality system of the minimization problem in the deterministic problem; e.g., [32]. In the stochastic case, however, we first need to show that there is a Lagrange multiplier before using the Lagrange multiplier rule to derive an optimality system of stochastic equations. To show the existence of a Lagrange multiplier, we follow the method given in [28].

We begin with the definition of the abstract class of minimization problems. Let G,X, and Y be reflexive Banach spaces whose norms are denoted by $\|\cdot\|_G$,$\|\cdot\|_X$, and $\|\cdot\|_Y$ and whose dual spaces are denoted by G^*,X^*, and Y^*, respectively. Let Θ be the control set that is a closed convex subset of G.

We assume that the functional to be minimized takes the form

$$\mathcal{J}(v,z)=\lambda\mathcal{F}(v)+\lambda\mathcal{E}(z) \quad \forall(v,z)\in X\times\Theta, \tag{4.11}$$

where $\mathcal{F}$ is a functional on X, $\mathcal{E}$ is a functional on Θ, and λ is a given parameter that is assumed to belong to a compact interval $\Lambda\subset\mathbb{R}_+$.

We define the function $M:X\times\Theta\to X$ for the constraint equation $M(v,z)=0$ as follows:

$$M(v,z)=v+\lambda TN(v)+\lambda TK(z) \quad \forall(v,z)\in X\times\Theta, \tag{4.12}$$

where $N:X\to Y$ is a differentiable map, $K:\Theta\to Y$ is a bounded linear operator, $T:Y\to X$ is a bounded linear operator, and $\lambda\in\Lambda$.

With these definitions, we now consider the following constrained minimization problem:

$$\min_{(v,z)\in X\times\Theta} \mathcal{J}(v,z) \quad \text{subject to} \quad M(v,z)=0. \tag{4.13}$$

The set of hypotheses needed to justify the use of the Lagrange multiplier rule and to derive an optimality system can be determined is given by

(HE1) for each $z\in\Theta$, $v\mapsto\mathcal{J}(v,z)$ and $v\mapsto M(v,z)$ are Fréchet differentiable;

(HE2) $z\mapsto\mathcal{E}(z)$ is convex;

(HE3) for $v\in X$, $N'(v)$ maps from X into $Z\hookrightarrow\hookrightarrow Y$, where N' denotes the Fréchet derivative of N.

Theorem 4.2. *Let $\lambda\in\Lambda$ be given. Assume that there exists an optimal solution (u,f) of* (4.13) *in $X\times\Theta$ and that* (HE1)-(HE3) *hold. Then there exists a $k\in\mathbb{R}$ and a $\mu\in X^*$ that are not both* *equal to zero such that*

$$\begin{aligned}&k\langle \mathcal{J}_u(u,f),w\rangle - \langle \mu, M_u(u,f)\cdot w\rangle = 0 \quad \forall w\in X,\\ &\min_{z\in\Theta}\mathcal{L}(u,z,\mu,k) = \mathcal{L}(u,f,\mu,k).\end{aligned}$$

Theorem 4.3. *Let $\lambda\in\Lambda$ be given. Assume that there exists an optimal solution (u,f) of* (4.13) *in $X\times G$, that* (HE1)-(HE3) *hold, and that the mapping $z\mapsto\mathcal{E}(z)$ is Fréchet differentiable on G. Then there exists a $k\in\mathbb{R}$ and a $\mu\in X^*$, not both equal to zero, such that*

$$\begin{aligned}&k\langle \mathcal{J}_u(u,f),w\rangle - \langle \mu,(I+\lambda TN'(u))\cdot w\rangle = 0 \quad &&\forall w\in X,\\ &k\langle \mathcal{E}'(f),z\rangle - \langle \mu, TKz\rangle = 0 &&\forall z\in G.\end{aligned}$$

Remark 4.1. For two Theorems 4.2 and 4.3, if $1/\lambda$ is not in $\sigma(-TN'(u))$, we may choose $k=1$; see [28].

4.3 The existence of Lagrange multipliers and the optimality system of stochastic equations

We are now ready to prove the existence of a Lagrange multiplier for our minimization problem (4.10). The Lagrange multiplier rule may be used to convert the constrained minimization problem into an unconstrained one. Then we find the optimality system of stochastic equations.

Note that since our stochastic elliptic PDE has a unique solution regardless of the choice of λ, a parameter in the abstract setting, we take $\lambda=1$.

Recall the stochastic optimal control problem:

$$\min \mathcal{J}(u,f) \quad \text{subject to} \quad M(u,f)=0 \quad \forall v\in\mathcal{H}_0^1(D), \tag{4.14}$$

where $M(u,f)=b[u,v]-[f,v]$.

We define $X=\mathcal{H}_0^1(D)$, $Y=\mathcal{H}^{-1}(D)$, $G=L^2(D)$, and $Z=\{0\}$. Then clearly we have $Z\hookrightarrow\hookrightarrow Y$. For the time being, we assume that the admissible set Θ for the control f is a closed, convex subset of G. We define the continuous linear operator $T\in\mathcal{L}(Y;X)$ as follows. For $g\in Y$, $Tg=u\in X$ is the unique solution of

$$b[u,v]=[g,v] \quad \forall v\in X. \tag{4.15}$$

We define the (differentiable) mapping $N:X\to Y$ by

$$\langle N(u),v\rangle = 0 \quad \forall v\in X \tag{4.16}$$

and define $K:G\to Y$ by

$$\langle Kf,\eta\rangle = -\langle f,\eta\rangle \quad \forall \eta\in X. \tag{4.17}$$

Then the constraint equation (2.8) can be expressed by $u+TKf=0$. We note that

$$\mathcal{F}(u)=\mathbb{E}\left(\frac{1}{2}\int_D|u-U|^2\,dx\right) \quad \text{and} \quad \mathcal{E}(f)=\mathbb{E}\left(\frac{\beta}{2}\int_D|f|^2\,dx\right). \tag{4.18}$$

Next, we verify the hypotheses for the existence of Lagrange multipliers. First, notice that (HE1) is obvious. Second, (HE2) holds because $f\mapsto\mathcal{E}(f)=\frac{\beta}{2}\|f\|^2_{\mathcal{L}^2(D)}$ is convex. Third, because for $\forall u\in X$, $N'(u)\cdot v=0\in Z\hookrightarrow\hookrightarrow Y$ for $\forall v\in X$, (HE3) holds.

The Lagrangian is given by

$$\mathcal{L}(u,f,\xi,k)=k\mathcal{J}(u,f)-b[u,\xi]+[f,\xi]$$

$\forall(u,f,\xi,k)\in X\times G\times X\times\mathbb{R}$.

By Theorem 4.2, there exists $\xi=T^*\mu\in X$ such that

$$\xi-kT^*\mathcal{F}'(u)=0, \tag{4.19}$$
$$\mathcal{L}(u,f,\xi,k)\le\mathcal{L}(u,z,\xi,k)\quad\forall z\in\Theta. \tag{4.20}$$

With $k=1$, (4.19) becomes

$$b[\xi,\zeta]=[u-U,\zeta]\quad\forall\zeta\in X \tag{4.21}$$

and (4.20) implies that

$$\frac{\beta}{2}[z,z]+[z,\xi]-\frac{\beta}{2}[f,f]+[f,\xi]\ge0\quad\forall z\in\Theta\subseteq G. \tag{4.22}$$

For each $\epsilon\in(0,1)$ and each $t\in\Theta$, set $z=\epsilon t+(1-\epsilon)f\in\Theta$. Then from (4.22), we have

$$\frac{\beta\epsilon}{2}[t-f,t-f]+\beta[t-f,f]+[t-f,\xi]\ge0\quad\forall t\in\Theta. \tag{4.23}$$

By letting $\epsilon\to0+$ in the above inequality, we have

$$[t-f,\beta f+\xi]\ge0\quad\forall t\in\Theta. \tag{4.24}$$

We now consider the case $\Theta=G$. Note that the mapping $z\mapsto\mathcal{E}(z)$ is Fréchet differentiable on G. Hence, by Theorem 4.3, (4.24) becomes an equality and by letting $z=t-f$ we obtain

$$[\beta f+\xi,z]=0\quad\forall z\in G. \tag{4.25}$$

The system formed by Eqs. (2.8), (4.21), and (4.25), which are necessary conditions for an optimum, is called an optimality system. We now conclude this section with the following theorem.

Theorem 4.4. *Let* $(u,f)\in\mathcal{H}^1_0(D)\times L^2(D)$ *be an optimal solution of* (4.10). *Then there exists* $\xi\in\mathcal{H}^1_0(D)$ *such that* (4.21) *and* (4.25) *hold.*

5 Discrete approximation of the optimality system

In this section, we solve stochastic optimal control problems using results from previous chapter and the Brezzi-Rappaz-Raviart theory. For this, we first introduce the theory. Throughout this section, we assume that $f\in L^2(D)$ for regularity of the solution.

5.1 Description of the Brezzi-Rappaz-Raviart theory

The B-R-R theory implies that the error of approximation of solutions of certain nonlinear problems under certain hypotheses is basically the same as the error of approximation of solutions of related linear problems; see [9,13,24]. Here for the sake of completeness, we will state the relevant results, specialized to our needs.

Consider the following type of nonlinear problems: seek $\psi\in\mathcal{X}$ such that

$$\psi+\mathcal{T}\mathcal{G}(\psi)=0, \tag{5.1}$$

where $\mathcal{T}\in\mathcal{L}(\mathcal{Y};\mathcal{X})$, $\mathcal{G}$ is a C^2 mapping from $\mathcal{X}$ into $\mathcal{Y}$, and $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces. We say that ψ is a regular solution of (5.1) if (5.1) holds and $\psi+\mathcal{T}\mathcal{G}_\psi(\psi)$ is an isomorphism from $\mathcal{X}$ into $\mathcal{X}$. Here $\mathcal{G}_\psi$ denotes the Fréchet derivative of $\mathcal{G}$ with respect to ψ. We assume that there exists another Banach space $\mathcal{Z}$, contained in $\mathcal{Y}$, with continuous imbedding, such that

$$\mathcal{G}_\psi(\psi)\in\mathcal{L}(\mathcal{X};\mathcal{Z})\quad\forall\psi\in\mathcal{X}. \tag{5.2}$$

Approximations are defined by introducing a subspace $\mathcal{X}^h\subset\mathcal{X}$ and an approximating operator $\mathcal{T}^h\in\mathcal{L}(\mathcal{Y};\mathcal{X}^h)$. We seek $\psi^h\in\mathcal{X}^h$ such that

$$\psi^h+\mathcal{T}^h\mathcal{G}(\psi^h)=0. \tag{5.3}$$

Concerning the operator $\mathcal{T}^h$, we assume the approximation properties

$$\lim_{h\to0}\|(\mathcal{T}^h-\mathcal{T})\omega\|_{\mathcal{X}}=0\quad\forall\omega\in\mathcal{Y} \tag{5.4}$$

and

$$\lim_{h\to0}\|\mathcal{T}^h-\mathcal{T}\|_{\mathcal{L}(\mathcal{Z};\mathcal{X})}=0. \tag{5.5}$$

Note that whenever the imbedding $\mathcal{Z}\subset\mathcal{Y}$ is compact, (5.5) follows from (5.4) and, moreover, (5.2) implies that the operator $\mathcal{T}\mathcal{G}_\psi(\psi)\in\mathcal{L}(\mathcal{X};\mathcal{X})$ is compact.

We now state the result of [9] that will be used in the sequel. In the statement of the theorem, $D^2\mathcal{G}$ represents any and all second Fréchet derivatives of $\mathcal{G}$.

Theorem 5.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. Assume that $\mathcal{G}$ is a C^2 mapping from $\mathcal{X}$ to $\mathcal{Y}$ and that $D^2\mathcal{G}$ is bounded on all bounded sets of $\mathcal{X}$. Assume that* (5.2), (5.4), *and* (5.5) *hold and that ψ is a regular solution of* (5.1). *Then there exists a neighborhood $\mathcal{O}$ of the origin in $\mathcal{X}$ and, for $h\le h_0$ small enough, a unique $\psi^h\in\mathcal{X}^h$ such that ψ^h is a regular solution of* (5.3). *Moreover, there exists a constant $C>0$, independent of h, such that*

$$\|\psi^h-\psi\|_{\mathcal{X}}\le C\|(\mathcal{T}^h-\mathcal{T})\mathcal{G}(\psi)\|_{\mathcal{X}}. \tag{5.6}$$

5.2 Recasting the optimality system and its discrete approximation into the B-R-R framework

We first fit our optimality system and its discrete approximation into the B-R-R framework to derive error estimates for the discrete approximation of the optimality system. Then we obtain the desired error estimates by verifying assumptions in the B-R-R theory.

We set $\mathcal{X}=\mathcal{H}_0^1(D)\times L^2(D)\times\mathcal{H}_0^1(D)$ and $\mathcal{Y}=H^{-1}(D)\times\mathcal{H}_0^1(D)$. We 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

$$b[\tilde{u},v]=[\tilde{r},v]\quad\forall v\in\mathcal{H}_0^1(D),\tag{5.7}$$
$$b[\tilde{\xi},\zeta]=[\tilde{\tau},\zeta]\quad\forall\zeta\in\mathcal{H}_0^1(D),\tag{5.8}$$
$$[\beta\tilde{f}+\tilde{\xi},z]=0\quad\forall z\in L^2(D).\tag{5.9}$$

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

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

Then it is clear that the optimality system (2.8), (4.21), and (4.25) can be written as

$$(u,f,\xi)+\mathcal{T}(\mathcal{G}(u,f,\xi))=0.\tag{5.10}$$

Hence, the optimality system is recast into the form of (5.1).

We now set $\mathcal{X}^{h\delta}=V^{h\delta}\times G^h\times V^{h\delta}$, where $V^{h\delta}$ and G^h are from Section 3.1.

We define the discrete operator $\mathcal{T}^{h\delta}\in\mathcal{L}(\mathcal{Y};\mathcal{X}^{h\delta})$ as follows:

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

if and only if

$$b[\tilde{u}^{h\delta},v^{h\delta}]=[\tilde{r},v^{h\delta}]\quad\forall v^{h\delta}\in V^{h\delta},\tag{5.11}$$
$$b[\tilde{\xi}^{h\delta},\zeta^{h\delta}]=[\tilde{\tau},\zeta^{h\delta}]\quad\forall\zeta^{h\delta}\in V^{h\delta},\tag{5.12}$$
$$[\beta\tilde{f}^h+\tilde{\xi}^{h\delta},z^h]=0\qquad\forall z^h\in G^h.\tag{5.13}$$

Then it is clear that the discrete optimality system,

$$b[u^{h\delta},v^{h\delta}]=[f^h,v^{h\delta}]\qquad\forall v^{h\delta}\in V^{h\delta},\tag{5.14}$$
$$b[\xi^{h\delta},\zeta^{h\delta}]=[u^{h\delta}-U,\zeta^{h\delta}]\quad\forall\zeta^{h\delta}\in V^{h\delta},\tag{5.15}$$
$$[\beta f^h+\xi^{h\delta},z^h]=0\qquad\forall z^h\in G^h,\tag{5.16}$$

can be written as

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

Hence, the discrete optimality system is recast into the form of (5.3).

5.3 Error estimates for the approximation of solutions of the optimality system

In this section, we proceed to verify all assumptions in Theorem 5.1. We define first a space $\mathcal{Z}=L^2(D)\times\mathcal{L}^2(D)$. Then clearly this space is continuously embedded into $\mathcal{Y}=H^{-1}(D)\times\mathcal{H}^{-1}(D)$.

Denote the Fréchet derivative of $\mathcal{G}(u,f,\xi)$ with respect to (u,f,ξ) by $D\mathcal{G}(u,f,\xi)$ or $\mathcal{G}_{(u,f,\xi)}(u,f,\xi)$. Then from $\mathcal{G}(u,f,\xi)$, 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 the stochastic optimal control problems.

Proposition 5.1.

1. *$D\mathcal{G}(u,f,\xi)\in\mathcal{L}(\mathcal{X};\mathcal{Z})$ for all $(u,f,\xi)\in\mathcal{X}$.*
2. $\mathcal{G}$ is twice continuously differentiable and $D^2\mathcal{G}$ is bounded on all bounded sets of $\mathcal{X}$.
3. For any $(\tilde{r},\tilde{\tau})\in\mathcal{Y}$, $\|(\mathcal{T}-\mathcal{T}^{h\delta})(\tilde{r},\tilde{\tau})\|_{\mathcal{X}}\to 0$ as $h,\delta\to 0$.
4. $\|\mathcal{T}-\mathcal{T}^{h\delta}\|_{\mathcal{L}(\mathcal{Z},\mathcal{X})}\to 0$ as $h,\delta\to 0$.
5. A solution of (5.10) is regular.

Proof. Similar proofs can be found in [29]. □

Note that the facts in Proposition 5.1 are all the assumptions of Theorem 5.1. Thus, by that theorem, we obtain the following results.

Theorem 5.2. *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* (2.8), (4.21), *and* (4.25). *Let $(u^{h\delta},f^h,\xi^{h\delta})\in V^{h\delta}\times G^h\times V^{h\delta}$ be the solution of the discrete optimality system* (5.14), (5.15), *and* (5.16). *Then we have*

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

Moreover, there exists $C>0$ such that

$$\begin{aligned}&\|u-u^{h\delta}\|^2_{\mathcal{H}_0^1(D)}+\|f-f^h\|^2_{L^2(D)}+\|\xi-\xi^{h\delta}\|^2_{\mathcal{H}_0^1(D)}\\&\le C(h^2+\delta^{2\gamma})K(\|f\|^2_{L^2(D)}+\|u-U\|^2_{\mathcal{L}^2(D)}),\end{aligned}\tag{5.17}$$

where

$$K=\max\left\{1,\frac{1}{(k!)^2}\ \|\phi_j\|^{2(p_j+1-k)}_{L^\infty(D)}\ :\ 1\le j\le N,\ 0\le k\le p_j+1\right\}.$$

Remark 5.1. In the above theorem, we used $\mathcal{H}_0^1(D)=L^2(\Gamma;H_0^1(D))$ instead of $C^{p+1}(\Gamma;H_0^1(D))$ since the assumption that our coefficient a has a finite K-L expansion guarantees that our solution u is in $C^{p+1}(\Gamma;H_0^1(D))$ (see Lemma 4.1 in [34] or Remark 5.1 in [5]). As a result, with the truncated K-L expansion of a, our numerical solution is supposed to be $p+1$ times differentiable in the y-direction.

6 Numerical computation of stochastic control problems

In this section, we consider the space $Z^p \subset L^2(D)$, where $Z^p = Z_1^{p_1} \times Z_2^{p_2} \times \cdots \times Z_N^{p_N}$ and $Z_n^{p_n} = \{v:\Gamma_n \to \mathbb{R}: v \in span(1, y_n, \cdots, y_n^{p_n})\}$. This space is a particular case of the space Y^δ in Section 3.1 with no partition of Γ (instead, we increase only the polynomial degree). We then think of finite element spaces on $D\times\Gamma$, say V^{hp}. Here, if $v \in V^{hp}$, $v \in$ span($\phi\psi$: $\phi \in X^h$ and $\psi(y) \in Z^p$).

6.1 The discrete system of equations

Let $\{\varphi_i(x)\}$ be a basis of the space $X^h \subset H_0^1(D)$ and let $\{\psi_j(y)\}$ be a basis of the space $Z^p \subset L^2(D)$. Then the solution of the discrete optimality system of equations is given by

$$u^{hp}(x,y) = \sum_{i,j} u_{ij}\varphi_i(x)\psi_j(y), \tag{6.1a}$$

$$\xi^{hp}(x,y) = \sum_{i,j} \xi_{ij}\varphi_i(x)\psi_j(y), \tag{6.1b}$$

$$f^h(x) = \sum_i f_i\varphi_i(x). \tag{6.1c}$$

Recall the discrete optimality system of equations:

$$-\int_\Gamma \rho \int_D a\nabla u^{hp}\cdot\nabla v^{hp} dxdy + \int_\Gamma \rho \int_D f^h v^{hp} dxdy = 0 \qquad \forall v^{hp} \in V^{hp}, \tag{6.2a}$$

$$\int_\Gamma \rho \int_D \xi^{hp}\eta^h dxdy + \beta \int_\Gamma \rho \int_D f^h \eta^h dxdy = 0 \qquad \forall \eta^h \in G^h, \tag{6.2b}$$

$$\int_\Gamma \rho \int_D u^{hp}\lambda^{hp} dxdy - \int_\Gamma \rho \int_D a\nabla \xi^{hp}\nabla\lambda^{hp} dxdy = \int_\Gamma \rho \int_D U\lambda^{hp} dxdy \quad \forall \lambda^{hp} \in V^{hp}. \tag{6.2c}$$

By substituting (6.1) into this system of equations (6.2), we have

$$\begin{aligned}
&\int_\Gamma \rho(y) \int_D a(x,y)\nabla u^{hp}(x,y)\nabla v^{hp}(x,y)\, dxdy \\
&\qquad = \sum_{i,j}\left(\int_\Gamma \rho(y)\psi_j(y)\psi_l(y)\int_D a(x,y)\nabla\varphi_i(x)\nabla\varphi_k(x)\, dxdy\right) u_{ij}, \\
&\int_\Gamma \rho(y)\int_D f^h(x)v^{hp}(x,y)\, dxdy = \sum_i \left(\int_\Gamma \rho(y)\psi_l(y)\int_D \varphi_i(x)\varphi_k(x)\, dxdy\right) f_i, \\
&\int_\Gamma \rho(y)\int_D \xi^{hp}(x,y)\eta^h(x)\, dxdy = \sum_{i,j}\left(\int_\Gamma \rho(y)\psi_j(y)\int_D \varphi_i(x)\varphi_k(x)\, dxdy\right)\xi_{ij}, \\
&\beta\int_\Gamma \rho(y)\int_D f^h(x)\eta^h(x)\, dxdy = \sum_i\left(\beta\int_\Gamma \rho(y)\int_D \varphi_i(x)\varphi_k(x)\, dxdy\right) f_i,
\end{aligned}$$

$$
\begin{aligned}
&\int_\Gamma \rho(y)\int_D u^{hp}(x,y)\lambda^{hp}(x,y)dxdy=\sum_{i,j}\left(\int_\Gamma \rho(y)\psi_j(y)\psi_l(y)\int_D \varphi_i(x)\varphi_k(x)dxdy\right)u_{ij},\\
&\int_\Gamma \rho(y)\int_D a(x,y)\nabla\xi^{hp}(x,y)\nabla\lambda^{hp}(x,y)\ dxdy\\
&\qquad=\sum_{i,j}\left(\int_\Gamma \rho(y)\psi_j(y)\psi_l(y)\int_D a(x,y)\nabla\varphi_i(x)\nabla\varphi_k(x)\ dxdy\right)\xi_{ij},\\
&\int_\Gamma \rho(y)\int_D U(x,y)\lambda^{hp}(x,y)\ dxdy=\int_\Gamma \rho(y)\int_D U(x,y)\varphi_k(x)\psi_l(y)\ dxdy
\end{aligned}
$$

for any test function $\varphi_k(x)\psi_l(y)$.

We now look at only the right hand side of the first equation. Note that for $\psi_j(y)\in Z^p$, we have $\psi_j(y)=\prod_{m=1}^N \psi_{jm}(y_m)$, where $\psi_{jm}:\Gamma_m\to\mathbb{R}$ is a basis function of Z^{p_m}. With the truncated K-L expansion of $a(x,y)$, we obtain

$$
\begin{aligned}
&\int_\Gamma \rho(y)\psi_j(y)\psi_l(y)\int_D a(x,y)\nabla\varphi_i(x)\nabla\varphi_k(x)\ dxdy\\
=&\int_\Gamma \rho(y)\psi_j(y)\psi_l(y)\int_D \left(\mathbb{E}a(x,y)+\sum_{n=1}^N\sqrt{\lambda_n}\phi_n(x)y_n\right)\nabla\varphi_i(x)\nabla\varphi_k(x)\ dxdy\\
=&K^0_{i,k}\int_\Gamma \rho(y)\psi_j(y)\psi_l(y)\ dy+\sum_{n=1}^N K^n_{i,k}\int_\Gamma y_n\rho(y)\psi_j(y)\psi_l(y)\ dy\\
=&K^0_{i,k}\int_\Gamma \prod_{m=1}^N \rho_m(y_m)\psi_{jm}(y_m)\psi_{lm}(y_m)\ dy\\
&\quad+\sum_{n=1}^N K^n_{i,k}\int_\Gamma y_n\prod_{m=1}^N \rho_m(y_m)\psi_{jm}(y_m)\psi_{lm}(y_m)\ dy,
\end{aligned}
$$

where

$$
\begin{aligned}
K^0_{i,k}&=\int_D \mathbb{E}a(x,y)\nabla\varphi_i(x)\nabla\varphi_k(x)\ dx,\\
K^n_{i,k}&=\int_D \sqrt{\lambda_n}\phi_n(x)\nabla\varphi_i(x)\nabla\varphi_k(x)\ dx.
\end{aligned}
$$

Likewise, we could get the other equations. Next, we solve the linear system to determine u_{ij},ξ_{ij}, and f_i that are coefficients of solutions of the discrete optimality system of equations.

6.2 Numerical setting

In our numerical experiments, we assume for simplicity in calculation 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 the assumptions $\mathbb{E}X_n=0$ and $\mathrm{Var}X_n=1$

in the K-L expansion give that each uniform density function $\rho(X_n)$ is $\frac{1}{2\sqrt{3}}$. We thus assume that the joint probability density function ρ of $(X_1,X_2,\cdots,X_N)$ in our numerical experiments is $\frac{1}{(2\sqrt{3})^N}$.

We consider now $C(x_1,x_2)=e^{-|x_1-x_2|}$ as a covariance function and solve the following eigenvalue problem:

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

Then we have

$$\begin{aligned}
\phi_n(x)&=\frac{1}{\sqrt{1+\frac{\sin(2v_n)}{2v_n}}}\cos(v_n x), && \text{if n is odd,}\\
\phi_n(x)&=\frac{1}{\sqrt{1-\frac{\sin(2w_n)}{2w_n}}}\sin(w_n x), && \text{if n is even,}\\
\lambda_n&=\frac{2}{v_n^2+1}, && \text{if n is odd,}\\
\lambda_n&=\frac{2}{w_n^2+1}, && \text{if n is even,}
\end{aligned}$$

where v_n is a solution of $1-v\tan(v)=0$ and w_n is a solution of $w+\tan(w)=0$; see [23]. Note that λ_n gets smaller as v_n or w_n gets larger; see Fig. 1. This actually means that we would be okay with the truncated K-L expansion of the coefficient a for getting the solution to our model problem.

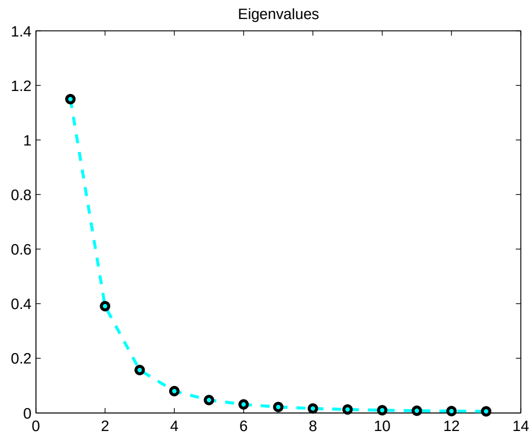


Figure 1: Eigenvalue decay.

6.3 Examples

We consider in this section a model problem under the homogeneous Dirichlet boundary condition with the target solution $U=\sin(\pi x)+\sin(2\pi x)$: Find the solution of

$$-(a(x,y)u'(x,y))'=f(x) \quad \forall(x,y)\in(-1,1)\times\prod_{n=1}^{N}(-\sqrt{3},\sqrt{3}), \tag{6.3a}$$

$$u(x,y)=0 \quad \forall(x,y)\in\{-1,1\}\times\prod_{n=1}^{N}(-\sqrt{3},\sqrt{3}) \tag{6.3b}$$

with flexible input data $f(x)$ to minimize

$$\mathcal{J}(u,f)=\frac{1}{2}\int_{-\sqrt{3}}^{\sqrt{3}}\frac{1}{(2\sqrt{3})^N}\int_{-1}^{1}|u-U|^2\,dxdy+\frac{\beta}{2}\int_{-1}^{1}|f|^2\,dx. \tag{6.4}$$

Note that here $'$ means differentiation with respect to x only and that the finite K-L expansion of $a(x,y)$ is given by

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

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

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

from previous section.

We first try to determine the value of an appropriate β. As we can see from both Fig. 2 and Table 1, for the fixed degrees of polynomials in each direction (in our first experiment, the maximum degree of polynomials in y_1-direction is 2 and y_2-direction is 1), we see that the expectation of the solution u to the optimality system of equations gets closer to the target solution $U=\sin(\pi x)+\sin(2\pi x)$ and $\mathbb{E}(\|u-U\|^2)$ and $\mathcal{J}(u,f)$ get smaller, respectively, as the value of β becomes smaller. From these outputs, we choose $\beta=10^{-8}$ for future use.

Table 1: $N=2$, $p=(2,1)$, $h=1/16$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$.

N	p	$\mathbb{E}(\|\|u-U\|\|^2)$	$\|\|f\|\|^2$	$\mathcal{J}(u,f)$	β
2	(2,1)	1.997425447584431	0.128653631686556	0.999355991950648	10^{-2}
2	(2,1)	1.780363586837130	1.039174048299372e+003	0.942140495833533	10^{-4}
2	(2,1)	0.335872718769967	3.144808537099188e+005	0.325176786239943	10^{-6}
2	(2,1)	0.002595035402397	1.385961737387668e+006	0.008227326388137	10^{-8}

In the second experiment (see Fig. 3 and Table 2), we see that for the fixed value of $\beta=10^{-8}$, the values of $\mathbb{E}(\|u-U\|^2)$ and $\mathcal{J}(u,f)$ are getting smaller as the value of a step size h gets smaller (these are expected results from our theory).

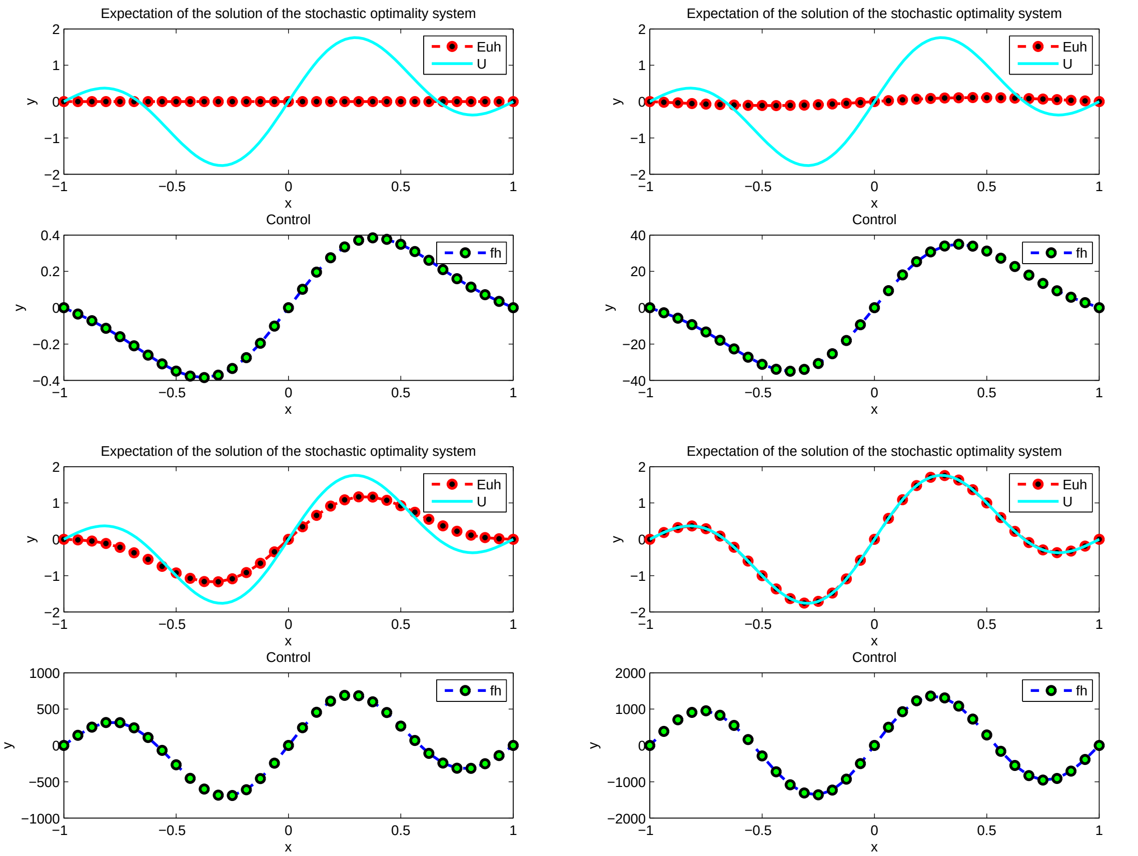


Figure 2: $N=2$, $p=(2,1)$, $h=1/16$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$, $\beta=10^{-2}$ (top left), $\beta=10^{-4}$ (top right), $\beta=10^{-6}$ (bottom left), $\beta=10^{-8}$ (bottom right).

Table 2: $N=2$, $p=(2,1)$, $\beta=10^{-8}$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$.

N	p	$\mathbb{E}(\|\|u-U\|\|^2)$	$\|\|f\|\|^2$	$\mathcal{J}(u,f)$	h
2	(2,1)	1.015620708248161	1.185634739035894e+005	0.508403171493598	1/2
2	(2,1)	0.017820040661929	1.920243067674646e+006	0.018511235669338	1/4
2	(2,1)	0.003235859956492	1.487385001888898e+006	0.009054854987691	1/8
2	(2,1)	0.002595035402397	1.385961737387668e+006	0.008227326388137	1/16
2	(2,1)	0.002553103337046	1.361403189137369e+006	0.008083567614210	1/32

We then use different polynomial degrees such as $p=(3,2,1)$ and $p=(4,2,2,1)$, which give us similar results as before (see Figs. 4 and 5, Tables 3 and 4).

Remark 6.1. In our paper, we have focused on the error analysis for a control tracking problem for SPDEs. As such, we have presented test examples to demonstrate both feasibility and efficiency of our theoretical convergence results and error estimates using

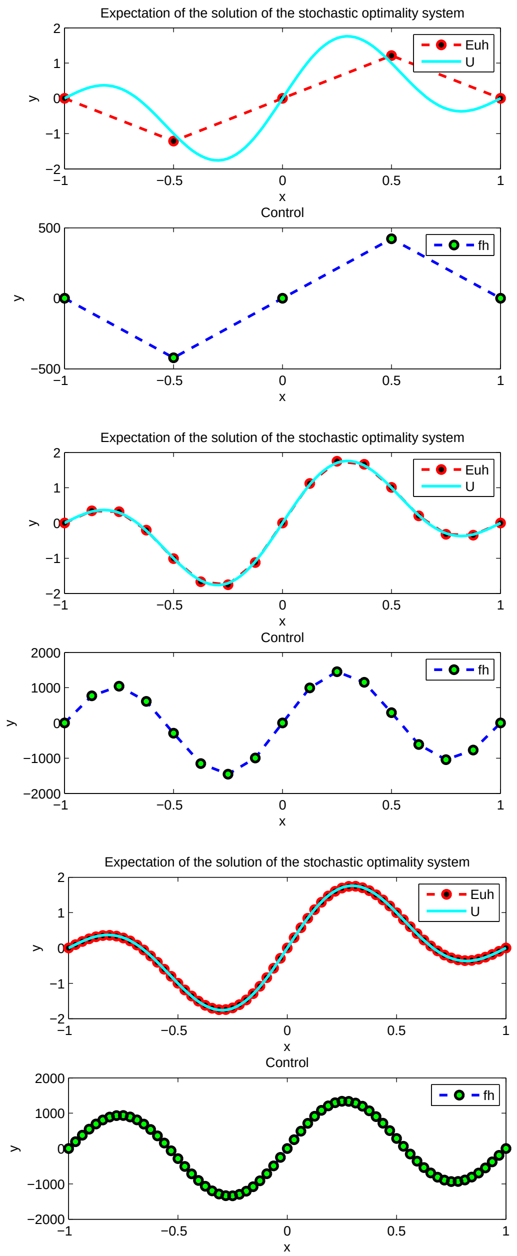
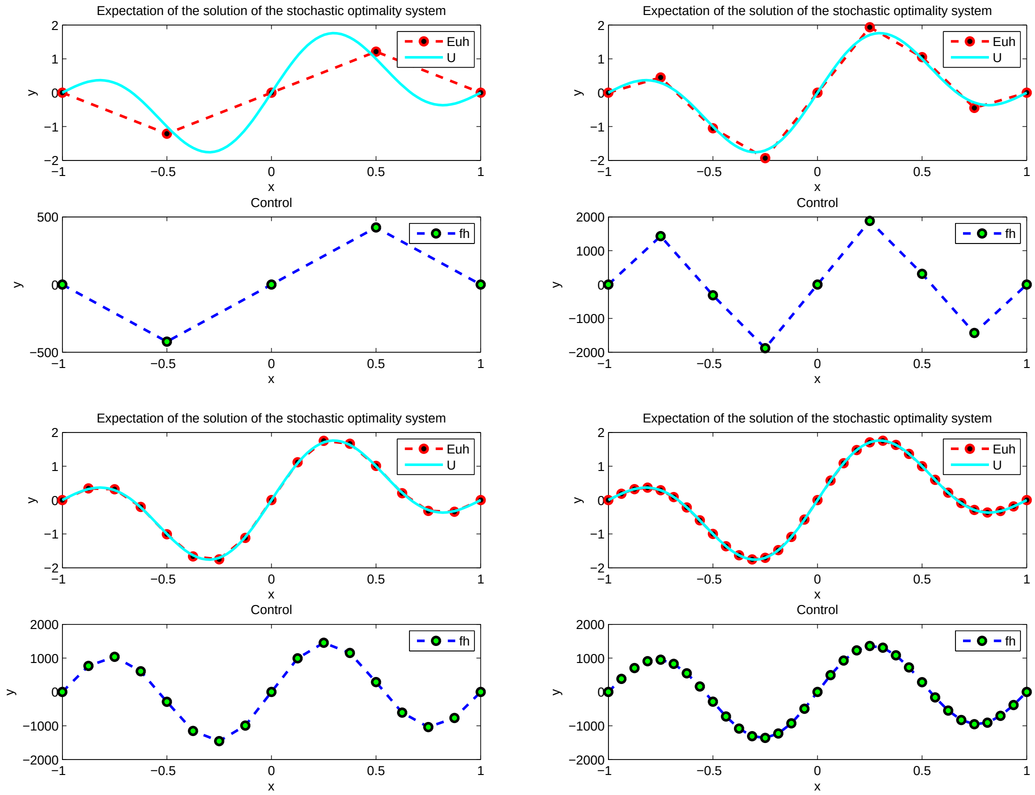


Figure 3: $N=2$, $p=(2,1)$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$, $h=1/2$ (top left), $h=1/4$ (top right), $h=1/8$ (middle left), $h=1/16$ (middle right), $h=1/32$ (bottom).

the solutions to Eq. (2.2) with the covariance function $e^{-|x_1-x_2|}$. As an ongoing project, we have been studying the case of a weakly correlated random field with a covariance function for which the analytical solution to Eq. (2.2) is not available. This and other numerical convergence and implementation issues including a white noise uncertainty will be addressed in a follow-up paper.



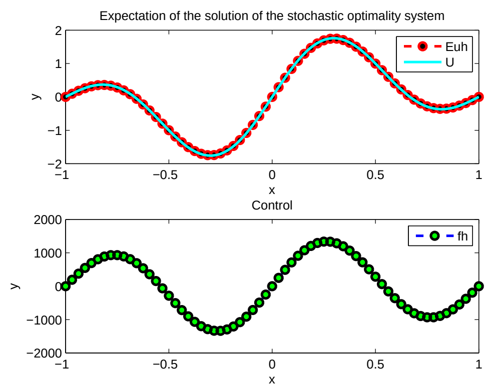


Figure 4: $N=3$, $p=(3,2,1)$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$, $h=1/2$ (top left), $h=1/4$ (top right), $h=1/8$ (middle left), $h=1/16$ (middle right), $h=1/32$ (bottom).

Table 3: $N=3$, $p=(3,2,1)$, $\beta=10^{-8}$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$.

N	p	$\mathbb{E}(\|\|u-U\|\|^2)$	$\|\|f\|\|^2$	$\mathcal{J}(u,f)$	h
3	(3,2,1)	1.015622289855443	1.185628036060636e+005	0.508403958945752	1/2
3	(3,2,1)	0.018016364044239	1.919216112496113e+006	0.018604262584600	1/4
3	(3,2,1)	0.003424910561098	1.486621685617537e+006	0.009145563708637	1/8
3	(3,2,1)	0.002782345885932	1.385255864189840e+006	0.008317452263915	1/16
3	(3,2,1)	0.002739990423212	1.360711073244207e+006	0.008173550577827	1/32

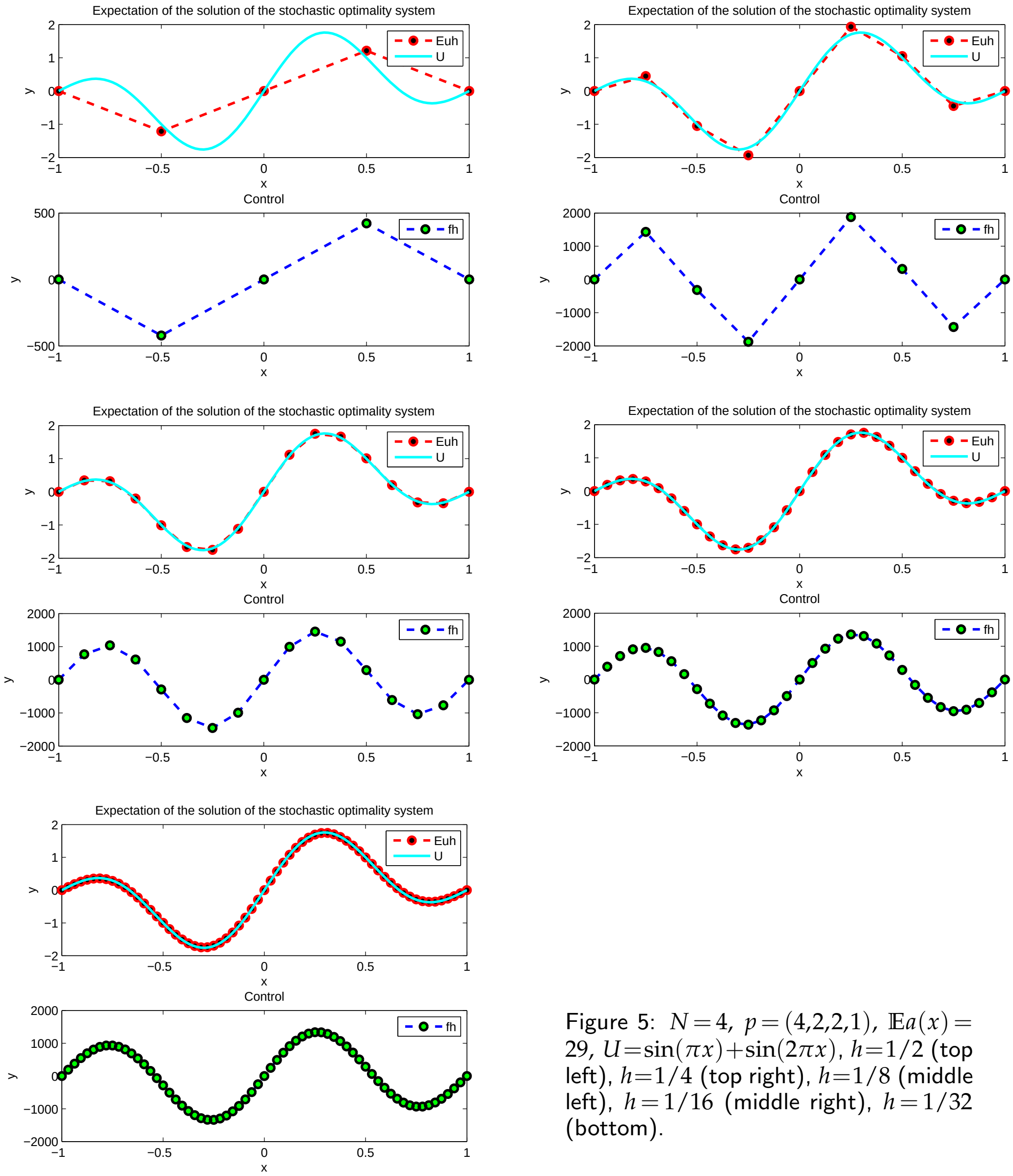


Figure 5: $N=4$, $p=(4,2,2,1)$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$, $h=1/2$ (top left), $h=1/4$ (top right), $h=1/8$ (middle left), $h=1/16$ (middle right), $h=1/32$ (bottom).

Table 4: $N=4$, $p=(4,2,2,1)$, $\beta=10^{-8}$, $\mathbb{E}a(x)=29$, $U=\sin(\pi x)+\sin(2\pi x)$.

N	p	$\mathbb{E}(\|\|u-U\|\|^2)$	$\|\|f\|\|^2$	$\mathcal{J}(u,f)$	h
4	(4,2,2,1)	1.015712744147771	1.185337096442278e+005	0.508449040622107	1/2
4	(4,2,2,1)	0.0181024011616080	1.918500726290845e+006	0.018643704212258	1/4
4	(4,2,2,1)	0.003500198538971	1.486120739084433e+006	0.009180702964907	1/8
4	(4,2,2,1)	0.002855149312878	1.384800188416728e+006	0.008351575598522	1/16
4	(4,2,2,1)	0.002812192486543	1.360266159504460e+006	0.008207427040794	1/32

Acknowledgments

The authors thank the anonymous referees for various useful comments that helped us improve the presentation of this paper. We would like to mention here that author Hyung-Chun Lee was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) (grant number NRF 2010-0026032); author Jangwoon Lee was supported by a faculty research grant from the University of Mary Washington.

References

[1] R. Adams, Sobolev Spaces, Academic, New York, 1975.
[2] I. Babuska and P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, Comput. Method Appl. Mech. Engrg. 191, pp. 4093–4122, 2002.
[3] I. Babuska, K. Liu, R. Tempone, Solving stochastic partial differential equations based on the experimental data, Math. Models Methods Appl. Sci. 13, no. 3, pp. 415–444, 2003.
[4] I. Babuska, R. Tempone, G. E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, Siam J. Numer. Anal. Vol. 42, no. 2, pp. 800–825, 2004.
[5] I. Babuska, R. Tempone, 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, pp. 1251–1294, 2005.
[6] S. Bonaccorsi, F. Confortola, and E. Mastrogiacomo, Optimal control of stochastic differential equations with dynamical boundary conditions, J. Math. Anal. Appl. 344, pp 667-681, 2008
[7] M. Bieri and C. Schwab, Sparse high order FEM for elliptic sPDEs, Comput. Methods Appl. Mech. Engrg., vol. 198, Issues 13-14, pp. 1149–1170, 2009.
[8] S. C. Brenner, L. R. Scott, The mathematical theory of finite element methods, Second Edition, Springer, 2002.
[9] F. Brezzi, J. Rappaz, P. Raviart, Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions, Numer. Math., 36 (1980), pp. 1-25
[10] N. J. Cutland and K. Grzesiak, Optimal Control for Two-Dimensional Stochastic Navier-Stokes Equations, Appl Math Optim 55:61–91, 2007.
[11] K. Chrysafinos, Moving mesh finite element methods for an optimal control problem for the advection-diffusion equation, Journal of Scientific Computing, Vol. 25, No. 3, pp:401–421, 2005.
[12] R. Courant and D. Hilbert, Methods of mathematical physics, Interscience, New York, 1953.
[13] M. Crouzeix, J. Rappaz, On numerical approximation in bifurcation theory, Masson, Parix, 1990.
[14] M. K. Deb, I. Babuska, J. T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Comput. Method Appl. Mech. Engrg. 190, pp. 6359–6372, 2001.
[15] A. Debussche, M. Fuhman, and G. Tessitore, Optimal Control of a Stochastic Heat Equation with Boundary-noise and Boundary-control,ESAIM:COCV, 13, pp 178–205, 2007.
[16] G. Da Prato, A. Debussche, Control of the Stochastic Burgers Model of Turbulent, SIAM J. Control Optim. Vol. 37, No. 4, pp. 1123–1149.

[17] A. Doostan, R. G. Ghanem, and J. Red-Horse, Stochastic model reduction for chaos representations, Comput. Methods Appl. Mech. Engrg., vol. 196, Issues 37-40, pp. 3951–3966, 2007.

[18] L. C. Evans, Partial differential equations, AMS, 1998.

[19] M. S. Gockenbach, Understanding and implementing the finite element method, SIAM, 2006.

[20] M. Loeve, Probability Theory Vol. 1 + 2, Springer, New York, 1978.

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

[22] J. Galvis and M. Sarkis, Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity, Siam J. Numer. Anal. Vol. 47, no. 5, pp. 3624–3651, 2009.

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

[24] V. Girault, P. Raviart, Finite element methods for Navier-Stokes equations, Springer, Berlin, 1986.

[25] M. S. Gockenbach, Understanding and implementing the finite element method, Siam, 2006.

[26] M. D. Gunzburger, L. S. Hou, 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, pp. 123–151, 1991.

[27] M. D. Gunzburger, L. S. Hou, 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, pp. 711–748, 1991.

[28] M. D. Gunzburger, L. S. Hou, Finite-dimensional approximation of a class of constrained nonlinear optimal control problems., SIAM J. Control. Optim., Vol. 34, no. 3, pp. 1001–1043, May 1996.

[29] M. D. Gunzburger, H.-C. Lee, and J. Lee, Error estimates of stochastic optimal Neumann boundary control problems., SIAM J. Numer. Anal. Vol. 49, No. 4, pp. 1532-1552, 2011.

[30] H. Holden, B. Øksendal, J. Ubøe, T. Zhang Stochastic partial differential equations; A modeling, white noise functional approach, Birkhäuser, 1996.

[31] L. S. Hou, S. S. Lavindran, A penalized neumann control approach for solving an optimal dirichlet control problem for the navier-stokes equations, SIAM J. Control. Optim., Vol. 36, no. 5, pp. 1795–1814, September 1998.

[32] L. S. Hou, J. Lee, A Robin-Robin non-overlapping Domain Decomposition Method for An Elliptic Boundary Control Problem, Int. J. Numer. Anal. Model. Vol. 8, No. 3, pp 443-465, 2011.

[33] 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, No. 1, pp 87-103, 2011.

[34] J. E. Lagness, General boundary value problems for differential equations of Sobolev type, SIAM J. Math. Anal., 3, pp. 105–119, 1972.

[35] J. L. Lions, E. Magenes, Non-homogeneous boundary value problems and applications Vol. 1, Springer, New York, 1972.

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

[37] H. Manouzi, A finite element approximation of linear stochastic PDEs driven by multiplicative white noise, Int. J. Comput. Math., Vol. 85, no. 3-4, pp. 527–546, 2008.

[38] H. Manouzi and L. S. Hou, An optimal control problem for stochastic linear PDE's driven

by a Gaussian white noise, Numerical Mathematics and Advanced Applications, Springer Berlin Heidelberg, pp. 629–636, 2008.

[39] E. Rosseel and G. N. Wells, Optimal control with stochastic PDE constraints and uncertain controls, Comput. Methods Appl. Mech. Engrg., Vol. 213-216, pp. 152–167, 2012.

[40] C. Schwab, R. A. Todor, Sparse finite elements for elliptic problems with stochastic loading, Numer. Math., pp. 707–734, 2003.

[41] G. Stefanou, The stochastic finite element method: Past, present, and future, Comput. Methods Appl. Mech. Engrg., vol. 198, Issues 9-12, pp. 1031–1051, 2009.