East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.041013.180314a

Vol. **4**, No. 2, pp. 166-188
May 2014

Sparse Grid Collocation Method for an Optimal Control Problem Involving a Stochastic Partial Differential Equation with Random Inputs

Nary Kim and Hyung-Chun Lee*

Department of Mathematics, Ajou University, Suwon, Korea 443-749.
Received 4 October 2013; Accepted (in revised version) 18 March 2014.
Available online 7 April 2014

Abstract. In this article, we propose and analyse a sparse grid collocation method to solve an optimal control problem involving an elliptic partial differential equation with random coefficients and forcing terms. The input data are assumed to be dependent on a finite number of random variables. We prove that an optimal solution exists, and derive an optimality system. A Galerkin approximation in physical space and a sparse grid collocation in the probability space is used. Error estimates for a fully discrete solution using an appropriate norm are provided, and we analyse the computational efficiency. Computational evidence complements the present theory, to show the effectiveness of our stochastic collocation method.

AMS subject classifications: 49A22, 49B22, 65M55, 65N30

Key words: Sparse grid collocation, stochastic partial differential equation, distributed control, finite element methods.

1. Introduction

We seek a solution (u,f) that minimises the cost functional

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

and satisfies the stochastic elliptic problem involving a Dirichlet boundary condition:

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

where $\mathbb{E}$ denotes expected value, D the spatial domain and ∂D its boundary, U a target solution to the constraint, β a positive constant influencing the relative importance of the two terms in Eq. (1.1), and f a stochastic control acting in the domain and depending on $a(x,\omega)$. The stochastic elliptic PDE generally models fluid flow in porous media; and under the homogeneous Dirichlet boundary condition, for almost every $\omega\in\Omega$ we look for a

*Corresponding author. *Email addresses:* `gyory@naver.com` (N. Kim), `hclee@ajou.ac.kr` (H.-C. Lee)

solution u that is a stochastic function from $\overline{D} \times \Omega$ to $\mathbb{R}$, where $D \subset \mathbb{R}^d$ is a convex bounded polygonal domain, and $a : D \times \Omega \to \mathbb{R}$ is a stochastic function. Here a has a bounded continuous covariance function in the context of a Karhunen-Loève (KL) expansion, with a uniformly bounded continuous first derivative, and ∇ means differentiation with respect to $x \in D$ only. This type of stochastic elliptic problem was previously investigated in the articles [2, 3, 12, 13, 22] and references therein.

To analyse this stochastic optimal control problem, we first estimate the error of the solution to the stochastic partial differential equation (SPDE), and then use the Brezzi-Rappaz-Raviart (BRR) theory to produce an error estimate of the solution to the stochastic optimal control problem. We then 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. In order to solve the stochastic optimal control problem numerically, we adopt a stochastic collocation method that has gained much attention recently in the computational community [4, 5]. Stochastic collocation can be based on either full or sparse tensor product approximation spaces, and seems to be ideal for computing statistics from solutions of PDE with random input data, since it essentially preserves the fast convergence of the spectral Galerkin method in maintaining an ensemble based approach (just as Monte Carlo). On the other hand, approximations based on tensor product grids suffer from the curse of dimensionality, since the number of collocations in a tensor grid grows exponentially fast in the number of input random variables. Thus even if the number of random variables is only moderately large, one should consider sparse tensor product spaces as first proposed by Smolyak [23]. Recently, total degree polynomial spaces and sparse tensor product spaces were investigated [8, 14, 24, 25]; and there have been substantial developments in stochastic collocation methods since [4, 5, 20, 21], where effective collocation strategies for problems involving a moderately large number of random variables have been devised.

The plan of this article is as follows. We represent a random field in Section 2, introducing the KL expansion and its truncated expansion. We also analyse our constraint equation and stochastic elliptic PDE, transforming a stochastic problem to a high-dimensional deterministic problem and presenting *a priori* error estimates. In Section 3, we introduce the discretisation method for probability space — viz. a sparse grid collocation method. In Section 4, the optimality system of equations is derived, showing the existence of a unique minimiser. We then establish our error estimate for the discrete approximate solutions to the optimality system, and in Section 5 give two numerical examples of stochastic optimal control problems constrained by the stochastic elliptic PDE under the Dirichlet boundary condition. Our brief concluding remarks are made in the final Section 6.

2. Preliminaries

2.1. Function spaces and problem setting

For our stochastic elliptic problem, we use a complete probability space $(\Omega, \mathscr{F}, P)$ where Ω is a set of outcomes, $\mathscr{F}$ is a σ-algebra of events, and $P : \mathscr{F} \to [0, 1]$ is a probability measure. We use standard Sobolev space notation [1]. For instance, $H^1(D)$ is a Hilbert space with norm $\|\cdot\|_{H^1(D)}$; and $H_0^1(D)$ is the subspace of $H^1(D)$ when the function value is zero on the boundary of D, with norm $\|u\|^2_{H_0^1(D)} = \int_D |\nabla u|^2\, dx$. Thus we define stochastic Sobolev spaces

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

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 $L^2(\Omega; L^2(D))$, and for simplicity let $\mathscr{L}^2(D) = L^2(\Omega; L^2(D))$ and $\mathscr{H}_0^1(D) = L^2(\Omega; H_0^1(D))$. These stochastic Sobolev spaces are Hilbert spaces.

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

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

and

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

where $\mathbb{E}$ denotes the expectation. Using (2.1) and (2.2), we can derive the weak formulation corresponding to the strong formulation (1.2) — viz. seek $u \in \mathscr{H}_0^1(D)$ such that

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

In this article, in order to ensure the existence and uniqueness of the solution to our stochastic elliptic problem (2.1), we assume that there are positive m and M such that

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

Then from the Lax-Milgram lemma [9], we have the following theorem.

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

$$b[u,v] = [f,v] \quad \forall\, v \in \mathscr{H}_0^1(D) \tag{2.5}$$

which satisfies the estimate

$$\|u\|_{\mathscr{H}_0^1(D)} \le \frac{C_P}{a_{min}} \left(\mathbb{E}\int_D |f|^2\, dx\right)^{\frac{1}{2}}, \tag{2.6}$$

where C_P follows from the Poincaré inequality

$$\|w\|_{L^2(D)} \le C_P \|\nabla w\|_{L^2(D)} \quad \forall\, w \in H_0^1(D)\,.$$

2.2. Karhunen-Loève expansions

We now introduce a Karhunen-Loève (KL) expansion, a well known theoretical tool for approximating stochastic functions [7, 14, 15, 19]. Thus if $a(x,\omega)$ is a stochastic function that has a continuous and bounded covariance function, it can be represented by

$$a(x,\omega)=\mathbb{E}a(x,\omega)+\sum_{n\geq 1}\sqrt{\lambda_n}\phi_n(x)X_n(\omega)\,, \tag{2.7}$$

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

$$\int_D C(x_1,x_2)\phi_n(x_1)\,dx_1=\lambda_n\phi_n(x_2)\,, \tag{2.8}$$

where $C(x_1,x_2)=\mathbb{E}(a(x_1,\omega)a(x_2,\omega))-\mathbb{E}a(x_1,\omega)\mathbb{E}a(x_2,\omega)$. This is the KL expansion of $a(x,\omega)$.

When we use numerical methods to approximate the solutions of mathematical models, we often use truncated expansions, and a random source in realistic mathematical models can be expressed by finitely many random variables. Consequently, in numerical methods we introduce truncations of the expansion (2.7) — i.e.

$$a_N(x,\omega)=\mathbb{E}a(x,\omega)+\sum_{n=1}^{N}\sqrt{\lambda_n}\phi_n(x)X_n(\omega). \tag{2.9}$$

However, the convergence of the truncated KL expansions is relevant to avoid modelling errors in the consequent numerical computations, and we have the following convergence theorem based on Mercer's theorem — e.g. see [6].

Theorem 2.1. *The truncated KL expansion* (2.9) *of a stochastic function* $a(x,\omega)$ *converges uniformly to* $a(x,\omega)$*:*

$$\sup_{x\in D}\mathbb{E}\left[(a(x,\omega)-a_N(x,\omega))^2\right]=\sup_{x\in D}\sum_{n=N+1}^{\infty}\lambda_n\phi_n^2(x)\rightarrow 0 \text{ as } N\rightarrow\infty\,. \tag{2.10}$$

In certain cases, we may need to ensure certain qualitative properties for the coefficiens a_N, and possibly describe them as nonlinear functions of X. The following standard transformation guarantees that the diffusivity coefficient is bounded away from zero a.s.:

$$\log(a_N-a_{min})(x,\omega)=b_0(x)+\sum_{n=1}^{N}\sqrt{\lambda_n}\phi_n(x)X_n(\omega) \tag{2.11}$$

— i.e. one performs a Karhunen-Loéve expansion for $\log(a_N-a_{min})$, assuming that $a>a_{min}$ a.s.. We will let $\Gamma_n\equiv X_n(\Omega)$ denote the image of X_n, $\Gamma=\prod_{n=1}^{N}\Gamma_n$, and assume that the random variables $[X_1,X_2,\cdots,X_N]$ have a joint probability density function $\rho:\Gamma\rightarrow\mathbb{R}_+$, with $\rho\in L^\infty(\Gamma)$.

The solution u of the constraint equation (1.2), a stochastic elliptic boundary value problem, can now be described by a finite number of random variables — i.e.

$$u_N(x,\omega) = u_N\left(x, X_1(\omega), \cdots, X_N(\omega)\right).$$

Under the above assumptions, we obtain the following high-dimensional deterministic equivalent weak formulation of (2.5) with the finite-dimensional information as follows: for given $f_N \in \mathscr{L}^2_\rho(D)$, find $u_N \in \mathscr{H}^1_{0,\rho}(D)$ such that

$$\int_\Gamma \rho(a_N \nabla u_N, \nabla v)_{L^2(D)}\, dy = \int_\Gamma \rho(f_N, v)_{L^2(D)}\, dy \quad \forall\, v \in \mathscr{H}^1_{0,\rho}(D). \tag{2.12}$$

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

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

where

$$\|v\|^2_{L^2_\rho(\Gamma; H^1_0(D))} = \int_\Gamma \rho\, \|v\|^2_{H^1_0(D)}\, dy = \mathbb{E}\|v\|^2_{H^1_0(D)}.$$

Similarly, we can define $L^2_\rho(\Gamma; L^2(D))$. For simplicity, we set $\mathscr{L}^2_\rho(D) = L^2_\rho(\Gamma; L^2(D))$ and $\mathscr{H}^1_{0,\rho}(D) = L^2_\rho(\Gamma; H^1_0(D))$ as before. The strong formulation of Eq. (2.12) is

$$\begin{aligned} &-\nabla \cdot \left[a_N(x,y)\nabla u_N(x,y)\right] = f(x,y) \quad \forall\, (x,y) \in D \times \Gamma, \\ &u_N(x,y) = 0 \quad \forall\, (x,y) \in \partial D \times \Gamma. \end{aligned} \tag{2.13}$$

Since our assumption (2.4) on $a(x,\omega)$ does not automatically imply the boundedness of the truncated KL expansion (2.9), in order 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 (2.4) — i.e. we assume that there exist $m, M > 0$ uniformly with respect to N such that

$$m \le \mathbb{E}a(x,\omega) + \sum_{n=1}^N \sqrt{\lambda_n}\phi_n(x)X_n(\omega) \le M \ \text{ a.e. for } (x,\omega) \in D \times \Omega, \tag{2.14}$$

when we have well-posedness of (2.13) because a_N is bounded. From the Lax-Milgram lemma [9], we have the following theorem about the existence and uniqueness of the solution u_N.

Theorem 2.2. *Let $f_N \in \mathscr{L}^2_\rho(D)$. Then there is a unique solution to the following weak formulation: find $u_N \in \mathscr{H}^1_{0,\rho}(D)$ such that*

$$b[u_N, v] = [f_N, v] \quad \forall v \in \mathscr{H}^1_{0,\rho}(D). \tag{2.15}$$

Moreover, u_N satisfies the estimate

$$\|u_N\|_{\mathscr{H}^1_0(D)} \le \frac{C_P}{m}\left(\mathbb{E}\int_D |f_N|^2\, dx\right)^{\frac{1}{2}}. \tag{2.16}$$

Now, we can estimate the modeling error produced by a perturbation of stochastic coefficients a and f of the problem (2.5) according to the following theorem — cf. Corollary 2.1 in Ref. [7] for a proof.

Theorem 2.3. *With the assumptions* (2.4), (2.14) *and* $f, f_N \in \mathscr{L}^2_\rho(D)$*, we have*

$$||u-u_N||_{\mathscr{H}^1_{0,\rho}(D)} \le \frac{C_P}{m}\left\{||f-f_N||_{\mathscr{L}^2_\rho(D)} + \frac{1}{m}||a-a_N||_{L^\infty_\rho(\Omega;L^\infty(D))}||f||_{\mathscr{L}^2_\rho(D)}\right\}. \tag{2.17}$$

We let $\zeta(N) := \frac{C_P}{m}\{||f-f_N||_{\mathscr{L}^2_\rho(D)} + \frac{1}{m}||a-a_N||_{L^\infty_\rho(\Omega;L^\infty(D))}||f||_{\mathscr{L}^2_\rho(D)}\}$, and assume that $\zeta(N)$ is a monotone decreasing function such that $\zeta(N) \to 0$ as $N \to \infty$.

3. Finite-Dimensional Approximations of PDE

We seek to approximate the exact solution of (1.2) in a suitable finite-dimensional subspace $V^{h,p}$, based on a tensor product $V^{h,p} = H_h(D) \otimes \mathscr{P}_{\mathbf{p}}(\Gamma)$. We introduce some standard approximate subspaces — viz. $H_h(D) \subset H^1_0(D)$ as a standard finite element space of dimension N_h, which contains continuous piecewise polynomials defined on regular triangulations $\mathscr{T}_h$ that have a maximum mesh-spacing parameter $h > 0$. Moreover, we suppose that H_h has the following approximation property: for a given function $u \in H^1_0(D)$,

$$\min_{v\in H_h(D)} ||u-v||_{H^1(D)} \le ch^s||u||_{H^{s+1}(D)}\,, \tag{3.1}$$

where s is a positive integer determined by the smoothness of u and the degree of the approximating finite element subspace, and c is independent of h. We also assume that there exists a finite element operator $\pi_h : H^1_0(D) \to H_h(D)$ with the optimally condition

$$||u-\pi_h u||_{H^1(D)} \le C_\pi \min_{v\in H_h(D)} ||u-v||_{H^1(D)} \le Ch^s||u||_{H^{s+1}(D)} \quad \forall\, u \in H^1_0(D)\,, \tag{3.2}$$

where the constant C_π is independent of the mesh size h and $C = cC_\pi$. For $\Gamma \subset \mathbb{R}^N$, $\mathscr{P}_{\mathbf{p}}(\Gamma) \subset \mathscr{L}^2(\Gamma)$ is the span of tensor product polynomials with degree at most $\mathbf{p} = (p_1,\dots,p_N)$ — i.e. $\mathscr{P}_{\mathbf{p}}(\Gamma) = \bigotimes_{n=1}^N \mathscr{P}_{p_n}(\Gamma_n)$, with

$$\mathscr{P}_{p_n}(\Gamma_n) = \mathrm{span}(y_n^k,\ k=0,\cdots,p_n)\,, \quad n=1,\cdots,N\,.$$

Stochastic collocation involves evaluation of approximations $\pi_h u_N(y_k) = u_h^N(y_k) \in H_h(D)$ to the solution u_N of (2.15) on a set of points $y_k \in \Gamma$, when the fully discrete solution $u^N_{h,\mathbf{p}} \in C^0(\Gamma; H_h(D))$ is a global approximation constructed from linear combinations of the point values. The approximation is thus given by

$$u^N_{h,\mathbf{p}}(y,x) = \sum_{k\in\mathscr{K}} u^N_h(y_k,x)l^{\mathbf{p}}_k(y)\,,$$

where the function $l_k^{\mathbf{p}}$ can be Lagrange polynomials. This formulation can be used to approximate the mean value or variance of u as

$$\mathbb{E}[u] \approx \mathbb{E}[u_h^N] \equiv \sum_k u_h^N(y_k,\cdot) \int_{\Gamma^N} l_k^{\mathbf{p}}(y)\rho(y)dy$$

and

$$\mathbb{V}ar[u] \approx \mathbb{V}ar[u_h^N] \equiv \sum_k \left(u_h^N(y_k,\cdot)\right)^2 \int_{\Gamma^N} l_k^{\mathbf{p}}(y)\rho(y)dy - \left(\bar{u}_h^N\right)^2.$$

3.1. Sparse collocation techniques

Sparse collocation methods have been investigated by many authors — cf. [4,5,20,21] and references therein. In multi-dimensional interpolation, one feasible methodology is to construct interpolants and nodal points by tensor products of one-dimensional interplant and nodal points, including Gauss and Chebyshev points that have been favoured since they both have low interpolation errors for polynomial approximation. An obvious disadvantage of this approach is that the number of points required increases combinatorially as the number of stochastic dimensions increased. The Smolyak algorithm provides one way to construct interpolation functions based on a minimal number of points in multi-dimensional space, where univariate interpolation formulas are extended to the multivariate case by using tensor products in a special way. This provides an interpolation strategy that potentially produces order of magnitude reductions in the number of support nodes required, and a linear combination of tensor products may be chosen such that the interpolation property is conserved for higher dimensions.

We introduce an index $i \in \mathbb{N}_+$, $i \geq 1$, and let $\{y_1^i, \cdots, y_{m_i}^i\} \subset [-1,1]$ be a sequence of abscissas for Lagrange interpolation on $[-1,1]$ for each value of i. For each direction y_n, we introduce a sequence of one-dimensional Lagrange interpolation operators of increasing order — viz. $\mathscr{U}^i : C^0(\Gamma_n; H_0^1(D)) \to V_{m_i}$ given by

$$\mathscr{U}^i(u)(y) = \sum_{j=1}^{m_i} u(y_j^i) l_j^i(y), \quad \forall\, u \in C^0\left(\Gamma_n; H_0^1(D)\right), \tag{3.3}$$

where

$$l_j^i(y) = \prod_{\binom{k=1}{k\neq j}}^{m_i} \frac{(y - y_k^i)}{(y_j^i - y_k^i)} \in \mathscr{P}_{m_i-1}(\Gamma_n)$$

are the Lagrange polynomials of degree $m-1$ and

$$V_{m_i}(\Gamma_n; H_0^1(D)) = \left\{ v \in C^0\left(\Gamma_n; H_0^1(D)\right) \ : \ v(y,x) = \sum_{k=1}^{m_i} \tilde{v}_k(x) l_k(y),\ \{\tilde{v}_k\}_{k=1}^{m_i} \in H_0^1 \right\}. \tag{3.4}$$

Formula (3.4) reproduces exactly all polynomials of degree less than m_i. Further, in the multivariate case $N > 1$, for each $u \in C^0(\Gamma; H_0^1(D))$ and the multi-index $\mathbf{i} = (i_1, \cdots, i_N) \in \mathbb{N}_+^N$ we define the full tensor product interpolation formulas

$$\mathscr{I}_{\mathbf{i}}^N u(y) = \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_N}\right)(u)(y) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_N=1}^{m_{i_N}} u\left(y_{j_1}^{i_1}, \cdots, y_{j_N}^{i_N}\right)\left(l_{j_1}^{i_1} \otimes \cdots \otimes l_{j_N}^{i_N}\right). \tag{3.5}$$

Clearly, the product involved above needs $\prod_{n=1}^N m_{i_n}$ function evaluations. These formulas will also be used as the building blocks for the Smolyak method, described next.

To analyse the convergence of various collocation methods, we note a regularity assumption on the data of the problem and consequent regularity results for the exact solution u_N. Thus we denote $\Gamma_n^* = \prod_{j=1, j\neq n}^N \Gamma_j$, let y_n^* be an arbitrary element of Γ_n^*, and make the following assumption concerning the solution to Eq. (2.12).

Assumption 3.1 (regularity)**.** For each $y_n \in \Gamma_n$, there exists $\tau_n > 0$ such that the function $u_N(x, y_n, y_n^*)$ as a function of $y_n, u_N : \Gamma_n \to C^0(\Gamma_n^*; H_0^1(D))$ admits an analytic extension $u(z, y_n^*, x)$, $z \in \mathbb{C}$ in the region of the complex plane

$$\Sigma(\Gamma_n; \tau_n) \equiv \left\{z \in \mathbb{C}, \mathrm{dist}(z, \Gamma_n)\right\}. \tag{3.6}$$

Moreover, $\forall\ z \in \Sigma(\Gamma_n; \tau_n)$

$$||u_N(z)||_{C^0(\Gamma*_n; H_0^1(D))} \leq \lambda, \tag{3.7}$$

with λ a constant independent of n.

It has been proven in Ref. [4] that problem (2.12) satisfies the analyticity results stated in Assumption 3.1. For instance, if we take $a(x, \omega)$ as the truncated expansion (2.11), then a suitable analyticity region $\Sigma(\Gamma_n; \tau_n)$ is given by

$$\tau_n = \frac{1}{4\sqrt{\lambda_n}||\phi_n||_{L^\infty(D)}}. \tag{3.8}$$

Note that since $\sqrt{\lambda_n}||\phi_n||_{L^\infty(D)} \to 0$ for a sufficiently regular covariance function, the analyticity region increases as n increases.

3.2. Smolyak approximation

Here we follow closely the work in Ref. [20] to describe the Smolyak *isotropic* formulas $\mathscr{A}(w, N)$. These formulas are just linear combinations of product formulas (3.5), with the following key properties where only products with a relatively small number of points are used. With $\mathscr{U}^0 = 0$, for $i \in \mathbb{N}_+$ we define

$$\Delta^i := \mathscr{U}^i - \mathscr{U}^{i-1}; \tag{3.9}$$

and given an integer $\mathrm{w} \in \mathbb{N}_+$, hereafter called the *level*, we define the sets

$$X(\mathrm{w},N) := \left\{ \mathbf{i} \in \mathbb{N}_+^N,\ \mathbf{i} \geq 1 \ :\ \sum_{n=1}^{N}(i_n - 1) \ \leq \ \mathrm{w} \right\} \tag{3.10}$$

$$\tilde{X}(\mathrm{w},N) := \left\{ \mathbf{i} \in \mathbb{N}_+^N,\ \mathbf{i} \geq 1 \ :\ \sum_{n=1}^{N}(i_n - 1) \ = \ \mathrm{w} \right\} \tag{3.11}$$

$$Y(\mathrm{w},N) := \left\{ \mathbf{i} \in \mathbb{N}_+^N,\ \mathbf{i} \geq 1 \ :\ \mathrm{w} - N + 1 \ \leq \ \sum_{n=1}^{N}(i_n - 1) \ \leq \ \mathrm{w} \right\} \tag{3.12}$$

and set $|\mathbf{i}| = i_1 + \cdots + i_N$ for $\mathbf{i} \in \mathbb{N}_+^N$. Then the isotropic Smolyak formula is given by

$$\mathscr{A}(w,N) = \sum_{\mathbf{i} \in X(w,N)} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_N}), \tag{3.13}$$

or equivalently

$$\mathscr{A}(w,N) = \sum_{\mathbf{i} \in Y(w,N)} (-1)^{\mathrm{w}+N-|\mathbf{i}|} \binom{N-1}{\mathrm{w}+N-|\mathbf{i}|} \cdot \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_N} \right). \tag{3.14}$$

Then in order to compute $\mathscr{A}(w,N)(u)$, one only needs to know function values on the "sparse grid"

$$\mathscr{H}(\mathrm{w},N) = \bigcup_{\mathbf{i} \in Y(\mathrm{w},N)} \left(\nu^{i_1} \times \cdots \times \nu^{i_N} \right) \subset [-1,1]^N, \tag{3.15}$$

where $\nu^i = \{y_1^i, \cdots, y_{m_i}^i\} \subset [-1,1]$ denotes the set of abscissas used by $\mathscr{U}^i$. If the sets are nested — i.e., $\nu^i \subset \nu^{i+1}$, then $\mathscr{H}(\mathrm{w},N) \subset \mathscr{H}(\mathrm{w}+1,N)$ and

$$\mathscr{H}(\mathrm{w},N) = \bigcup_{\mathbf{i} \in \tilde{X}(\mathrm{w},N)} \left(\nu^{i_1} \times \cdots \times \nu^{i_N} \right). \tag{3.16}$$

The Smolyak formula is actually interpolatory whenever nested points are used — cf. Ref. 8, Proposition 6 on the page 277.

On comparing (3.16) and (3.15), we observe that the Smolyak approximation that employs nested points requires fewer function evaluations than the corresponding formula with non-nested points. We introduce two particular sets of abscissas below, respectively nested and non nested.

3.3. Choice of interpolation abscissas

Clenshaw-Curtis abscissas. We first use Clenshaw-Curtis abscissas in the construction of the Smolyak formula, which are the extrema of Chebyshev polynomials, and for any choice of $m_i > 1$ are given by

$$y_j^i = -\cos\frac{\pi(j-1)}{m_i - 1}, \quad j = 1, \cdots, m_i. \tag{3.17}$$

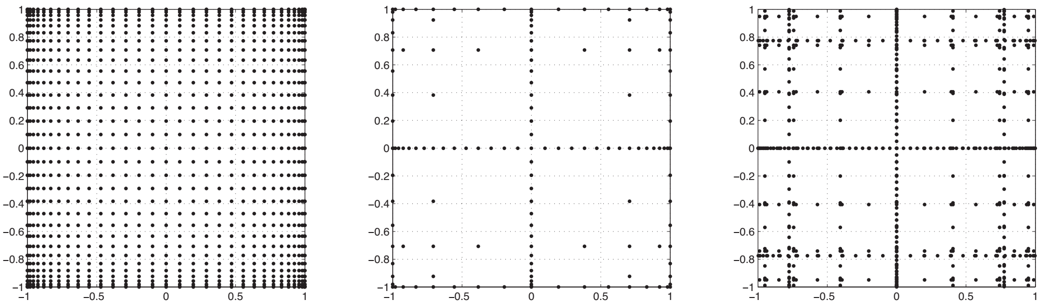

Figure 1: Full tensor product using Clenshaw-Curtis abscissas (left), Clenshaw-Curtis abscissas with level 5 (middle), and Gauss-Legendre abscissas with level 5 (right).

In addition, one sets $y_j^i = 0$ if $m_i = 1$ and lets the number of abscissas m_i in each level grow according to the formula

$$m_1 = 1 \qquad and \qquad m_i = 2^{i-1} + 1 \quad for \quad i > 1\,. \tag{3.18}$$

With this particular choice, one obtains nested sets of abscissas $\nu^i \subset \nu^{i+1}$, and thereby $\mathscr{H}(\mathrm{w},N) \subset \mathscr{H}(\mathrm{w}+1,N)$. It is important to choose $m_1 = 1$ if we are interested in the optimal approximation in relatively large N, because in all other cases the number of points used by $\mathscr{A}(\mathrm{w},N)$ increases too rapidly with N.

Gaussian abscissas. We also use Gaussian abscissas — i.e. the zeros of the orthogonal polynomials with respect to some positive weight. Although Gaussian abscissas are in general not nested, we choose the same number m_i of abscissas as in (3.18) in the Clenshaw-Curtis case. The natural choice of the weight should be the probability density function ρ of the random variables $Y_i(\omega), i = 1,\cdots,N$, but in the general multivariate case, the density ρ does not factorize (i.e. $\rho(y_1,\cdots,y_N) \neq \prod_{n=1}^N \rho_n(y_n)$) if the random variables Y_i are not independent. We therefore first introduce an auxiliary probability density function $\hat{\rho} : \Gamma \to \mathbb{R}^+$ that can be viewed as the joint probability of N independent random variable — i.e. it factorises as

$$\hat{\rho}(y_1,\cdots,y_N) = \prod_{n=1}^{N} \hat{\rho}_n(y_n) \quad \forall y \in \Gamma \tag{3.19}$$

and is such that $\left\|\rho/\hat{\rho}\right\|_{L^\infty(\Gamma)} < \infty$. For eaxh dimension $n = 1,\cdots,N$, we let the m_n Gaussian abscissas be the roots of the m_n degree polynomial that is $\hat{\rho}-$orthogonal to all polynomials of degree $m_n - 1$ on the interval $[-1,1]$. The auxiliary density $\hat{\rho}$ should be chosen as close as possible to the true density ρ, so that the quotient $\rho/\hat{\rho}$ is not too large (such a quotient will appear in the final error estimate).

3.4. Convergence analysis

The convergence analysis for the isotropic Smolyak method follows from Refs. [5,20]. A general approach is to bound both the discretisation error and the number of collocation points $\eta = \eta(w,N)$ in terms of the level w.

Collocation methods can be used to approximate the solution $u_N \in C^0(\Gamma^N; H_0^1(D))$ using finitely many function values, each computed by finite elements. Recalling that under Assumption 3.1 the $\{u_N\}$ admits an analytic extension, we let the fully discrete numerical approximation be $\mathscr{A}(w,N)\pi_h u_N$. Our aim is to give a priori estimates for the total error

$$e = u - \mathscr{A}(w,N)\pi_h u_N \,,$$

where the operator $\mathscr{A}(w,N)$ is described by Eq. (3.13) and π_h is the finite element projection operator. We investigate the error

$$||u - \mathscr{A}(w,N)\pi_h u_N|| \le ||u - u_N|| + ||u_N - \pi_h u_N|| + ||\pi_h u_N - \mathscr{A}(w,N)\pi_h u_N|| \,, \tag{3.20}$$

evaluated in the norm $L^q_\rho(\Omega; H_0^1(D))$ with either $q = 2$ or $q = \infty$.

To complete the convergence analysis, we introduce some lemmas from Ref. [5]. To do so, we introduce a weight $\sigma(y) = \prod_{n=1}^N \sigma_n(y_n) \le 1$ where

$$\sigma_n(y_n) = \begin{cases} 1\,, & \text{if } \Gamma_n \text{ is bounded}\,, \\ e^{-\alpha_n |y_n|} \text{ for some } \alpha_n > 0\,, & \text{if } \Gamma_n \text{ is unbounded}\,, \end{cases} \tag{3.21}$$

and the functional space

$$C^0_\sigma(\Gamma; V) \equiv \left\{ v : \Gamma \to V, v \text{ continuous in } y\,, \max_{y\in\Gamma} ||\sigma(y)v(y)||_V < +\infty \right\},$$

where V is a Banach space of functions defined in D.

Lemma 3.1. *If $f \in C^0_\sigma(\Gamma; L^2(D))$ and $a \in C^0_{loc}(\Gamma; L^\infty(D))$, uniformly bounded away from zero, then the solution to problem* (2.5) *satisfies $u \in C^0_\sigma(\Gamma; H_0^1(D))$.*

Lemma 3.2. *Under the assumption that for every $y = (y_n, y_n^*) \in \Gamma$ there exists $\gamma_n < +\infty$ such that*

$$\left\| \frac{\partial^k_{y_n} a(y)}{a(y)} \right\|_{L^\infty(D)} \le \gamma_n^k\, k! \quad \textit{and} \quad \frac{||\partial^k_{y_n} f(y)||_{L^2(D)}}{1 + ||f(y)||_{L^2(D)}} \le \gamma_n^k\, k!\,, \tag{3.22}$$

the solution $u(y_n, y_n^*, x)$ as a function of y_n, $u : \Gamma_n \to C^0_{\sigma_n^*}(\Gamma_n^*; H_0^1(D))$ admits an analytic extension $u(z, y_n^*, x)$, $z \in \mathbb{C}$, in the region of the complex plane

$$\Sigma(\Gamma_n; \tau_n) \equiv \{z \in \mathbb{C}\,,\ dist(z, \Gamma_n) \le \tau_n\} \tag{3.23}$$

with $0 < \tau_n < 1/(2\gamma_n)$. Moreover, for all $z \in \Sigma(\Gamma_n; \tau_n)$

$$||\sigma_n(Re\, z)u(z)||_{C^0_{\sigma_n^*}(\Gamma_n^*; H_0^1(D))} \le \frac{C_P\, e^{\alpha_n \tau_n}}{m(1 - 2\tau_n\gamma_n)} \left(2||f||_{C^0_\sigma(\Gamma; H_0^1(D))} + 1 \right)\,, \tag{3.24}$$

with the constant C_P as in inequality (2.6).

We also have a convergence theorem from Ref. [5], which focuses on the sparse stochastic collocation discretisation error.

Theorem 3.2. *For a function $u \in C^0(\Gamma; W(D))$ satisfying the assumptions of Lemma* 3.2*, the isotropic Smolyak formula based on $\hat{\rho}$-Gaussian abscissas satisfies*

$$\begin{aligned} &||u - \mathscr{A}(w,N)u||_{L^2_\rho(\Gamma;V)} \leq \sqrt{||\rho/\hat{\rho}||_{L^\infty(\Gamma)}} C(r_{min},N)\eta^{-\mu}, \\ &\mu := \frac{r_{min}\ e\ \log(2)}{\zeta + \log(N)}\ , \end{aligned} \tag{3.25}$$

with $\zeta := 1 + (1 + \log_2(1.5))\log(2) \approx 2.1$.

The constant $C(r_{min}, N)$ is defined in formula (3.31) on page 2331 of Ref. [20], and tends to zero as $r_{min} \to \infty$.

Similar results for the isotropic Smolyak formula based on Clenshaw-Curtis abscissas can also be found in Ref. [20].

4. Stochastic Distributed Control Problems

4.1. Optimality solution of stochastic equations

Let us now examine the existence of an optimal solution that minimises our functional (1.1). We first define

$$\mathscr{U}_{ad} = \left\{(u,f) \in \mathscr{H}_0^1 \times \mathscr{L}^2 \text{ such that (2.5) satisfied and } \mathscr{J}(u,f) < \infty\right\} \tag{4.1}$$

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

$$\mathscr{J}(\hat{u}, \hat{f}) \leq \mathscr{J}(u,f)\,. \tag{4.2}$$

The constrained minimisation problem can then be written

$$\min_{(u,f)\in\mathscr{U}_{ad}} \mathscr{J}(u,f)\,, \ \text{ subject to } \ (2.5)\,. \tag{4.3}$$

Assume that $\hat{f} \in \mathscr{L}^2(D)$ is a minimiser of $\mathscr{J}$ and $\hat{u}$ is the corresponding state variable, and define the adjoint variable $\hat{\xi} \in \mathscr{H}_0^1(D)$ such that

$$b[\hat{\xi}, \zeta] = [\hat{u} - u_d, \zeta]\,, \quad \forall\ \zeta \in \mathscr{H}_0^1(D)\,. \tag{4.4}$$

Theorem 4.1. *$\mathscr{J}$ has a unique minimiser $\hat{f} \in \mathscr{L}^2(D)$ and it is determined by*

$$[\delta \hat{f}, g] = -[\hat{\xi}, g]\,, \quad \forall\ g \in \mathscr{L}^2(D)\,. \tag{4.5}$$

Proof. The existence and uniqueness follow from the standard theory of optimal controls [18]. Now for $f \in \mathscr{L}^2(D)$, let $u(f)$ be the solution of (2.3). Then $\forall\ g \in \mathscr{L}^2(D)$,

$$\frac{\partial}{\partial \epsilon} u(\hat{f} + \epsilon g)\Big|_{\epsilon=0} = u(g)\,.$$

Since $\hat{f}$ is a minimiser of $\mathscr{J}$,

$$\frac{d\mathscr{J}(\hat{u}(\hat{f}), \hat{f})}{df} = \frac{d}{d\epsilon}\mathscr{J}(\hat{u}(\hat{f} + \epsilon g), \hat{f} + \epsilon g)\Big|_{\epsilon=0} = 0\,,$$

which implies

$$\mathbb{E}\left[\int_D (\hat{u} - u_d) u(g)\, dx + \delta \int_D f g\, dx\right] = 0\,. \tag{4.6}$$

From (4.4) and integration by parts, we have

$$\begin{aligned}
\mathbb{E}\left[\int_D (\hat{u} - u_d) u(g)\, dx\right] &= \mathbb{E}\left[\int_D (\nabla \cdot (a \nabla \hat{\xi})) u(g)\, dx\right] \\
&= -\mathbb{E}\left[\int_{\partial D} (a \nabla \hat{\xi} \cdot \vec{n}) u(g)\, dx + \int_D a \nabla \hat{\xi} \nabla u(g)\, dx\right] \\
&= \mathbb{E}\left[\int_D \hat{\xi} \nabla \cdot (a \nabla u(g))\, dx\right] \\
&= \mathbb{E}\left[\int_D \hat{\xi} g\, dx\right]\,.
\end{aligned} \tag{4.7}$$

From (4.6) and (4.7), we obtain (4.5). □

From the above theorem , we conclude that solving the minimisation problem (4.2) is equivalent to solving the following optimality system of equations:

$$\begin{aligned}
&b[u, v] = [f, v]\,, \quad \forall\ v \in \mathscr{H}_0^1(D)\,, \\
&b[\xi, \zeta] = [u - u_d, \zeta]\,, \quad \forall\ \zeta \in \mathscr{H}_0^1(D)\,, \\
&\delta [f, g]_{\mathscr{L}^2(D)} = [\xi, g]_{\mathscr{L}^2(D)}\,, \quad \forall\ g \in \mathscr{L}^2(D)\,.
\end{aligned} \tag{4.8}$$

Thus the reduced optimality system becomes

$$\begin{aligned}
&b[u, v] = \left[-\frac{\xi}{\delta}, v\right]\,, \quad \forall\ v \in \mathscr{H}_0^1(D)\,, \\
&b[\xi, \zeta] = [u - u_d, \zeta]\,, \quad \forall\ \zeta \in \mathscr{H}_0^1(D)\,.
\end{aligned} \tag{4.9}$$

4.2. Recasting the optimality system and its discrete approximation into the BRR framework

We first fit the optimality system and its discrete approximation into the BRR framework to derive error estimates for the discrete approximation of the optimality system, and then obtain the desired error estimates by verifying assumptions in the BRR theory. The BRR theory implies that the error of approximation of solutions of some nonlinear problems under certain hypotheses is basically the same as the error of approximation of solutions of related linear problems [10, 11, 16]. For the sake of completeness, we state the relevant results, specialised to our needs.

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

$$\psi + \mathscr{T}\mathscr{G}(\psi) = 0\,, \tag{4.10}$$

where $\mathscr{T} \in \mathscr{L}(\mathscr{Y};\mathscr{X})$, $\mathscr{G}$ is a C^2 mapping from $\mathscr{X}$ into $\mathscr{Y}$, and $\mathscr{X}$ and $\mathscr{Y}$ are Banach spaces. We say that ψ is a regular solution of (4.10) if $\psi + \mathscr{T}\mathscr{G}_\psi(\psi)$ is an isomorphism from $\mathscr{X}$ into $\mathscr{X}$, where $\mathscr{G}_\psi$ denotes the Fréchet derivative of $\mathscr{G}$ with respect to ψ. We assume that there exists another Banach space $\mathscr{Z}$, contained in $\mathscr{Y}$, with continuous imbedding such that

$$\mathscr{G}_\psi(\psi) \in \mathscr{L}(\mathscr{X};\mathscr{Z})\,, \quad \forall\ \psi \in \mathscr{X}. \tag{4.11}$$

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

$$\psi^h + \mathscr{T}^h\mathscr{G}(\psi^h) = 0\,. \tag{4.12}$$

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

$$\lim_{h\to 0} \|(\mathscr{T}^h - \mathscr{T})\omega\|_{\mathscr{X}} = 0\,, \quad \forall\ \omega \in \mathscr{Y} \tag{4.13}$$

and

$$\lim_{h\to 0} \|\mathscr{T}^h - \mathscr{T}\|_{\mathscr{L}(\mathscr{Z};\mathscr{X})} = 0\,. \tag{4.14}$$

Note that whenever the imbedding $\mathscr{Z} \subset \mathscr{Y}$ is compact, Eq. (4.14) follows from (4.13) and moreover, (4.11) implies that the operator $\mathscr{T}\mathscr{G}_\psi(\psi) \in \mathscr{L}(\mathscr{X};\mathscr{X})$ is compact.

We now state the result of Ref. [10] to be used below. In the statement of the theorem, $D^2\mathscr{G}$ represents any and all second Fréchet derivatives of $\mathscr{G}$.

Theorem 4.2. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces. Assume that $\mathscr{G}$ is a C^2 mapping from $\mathscr{X}$ to $\mathscr{Y}$ and that $D^2\mathscr{G}$ is bounded on all bounded sets of $\mathscr{X}$. Assume that* (4.11), (4.13), *and* (4.14) *hold and that ψ is a regular solution of Eq.* (4.10). *Then there exists a neighborhood $\mathscr{O}$ of the origin in $\mathscr{X}$ and, for $h \le h_0$ small enough, a unique $\psi^h \in \mathscr{X}^h$ such that ψ^h is a regular solution of* (4.12). *Moreover, there exists a constant $C > 0$, independent of h, such that*

$$\|\psi^h - \psi\|_{\mathscr{X}} \le C\|(\mathscr{T}^h - \mathscr{T})\mathscr{G}(\psi)\|_{\mathscr{X}}\,. \tag{4.15}$$

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

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

if and only if

$$b[\tilde{u}, v] = [\tilde{r}, v] \quad \forall v \in \mathscr{H}_0^1(D)\,, \tag{4.16}$$

$$b[\tilde{\xi}, \zeta] = [\tilde{\tau}, \zeta] \quad \forall \zeta \in \mathscr{H}_0^1(D)\,, \tag{4.17}$$

and

$$[\beta \tilde{f} + \tilde{\xi}, z] = 0 \quad \forall z \in \mathscr{L}^2(D)\,. \tag{4.18}$$

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

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

Then it is clear that the optimality system (4.8) can be written as

$$(u, f, \xi) + \mathscr{T}(\mathscr{G}(u, f, \xi)) = 0\,, \tag{4.19}$$

hence the optimality system is recast into the form of (4.10).

We now set $\mathscr{X}^{hp} = V^{hp} \times G^{hp} \times V^{hp}$ and define the discrete operator $\mathscr{T}^{hp} \in \mathscr{L}(\mathscr{Y};\mathscr{X}^{hp})$ as follows:

$$(\tilde{u}^{hp}, \tilde{f}^{hp}, \tilde{\xi}^{hp}) = \mathscr{T}^{hp}(\tilde{r}, \tilde{\tau})$$

if and only if

$$b[\tilde{u}^{hp}, v^{hp}] = [\tilde{r}, v^{hp}] \quad \forall v^{hp} \in V^{hp}, \tag{4.20}$$

$$b[\tilde{\xi}^{hp}, \zeta^{hp}] = [\tilde{\tau}, \zeta^{hp}] \quad \forall \zeta^{hp} \in V^{hp}, \tag{4.21}$$

and

$$[\beta \tilde{f}^{h} + \tilde{\xi}^{hp}, z^h] = 0 \quad \forall z^h \in G^h. \tag{4.22}$$

Then it is clear that the discrete optimality system

$$b[u^{hp}, v^{hp}] = [-\xi^{hp}, v^{hp}] \quad \forall v^{hp} \in V^{hp}\,, \tag{4.23}$$

$$b[\xi^{hp}, \zeta^{hp}] = [u^{hp} - U, \zeta^{hp}] \quad \forall \zeta^{hp} \in V^{hp} \tag{4.24}$$

can be written as

$$(u^{hp}, \xi^{hp}) + \mathscr{T}^{hp}\left(\mathscr{G}(u^{hp}, \xi^{hp})\right) = 0\,,$$

hence the discrete optimality system is recast into the form of (4.12).

4.3. Error estimates for approximation of solutions of the optimality system

We proceed to verify all assumptions in Theorem 4.2. We first define a space $\mathscr{X} = \mathscr{L}^2(D) \times \mathscr{L}^2(D)$, clearly continuously embedded into $\mathscr{Y} = \mathscr{H}^{-1}(D) \times \mathscr{H}^{-1}(D)$. If the Fréchet derivative of $\mathscr{G}(u,f,\xi)$ with respect to (u,f,ξ) us denoted by $D\mathscr{G}(u,f,\xi)$ or $\mathscr{G}_{(u,f,\xi)}$ (u,f,ξ), then from $\mathscr{G}(u,f,\xi)$ for $(u,f,\xi) \in \mathscr{X}$ we obtain

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

There are now the following propositions leading to our error analysis for the stochastic optimal control problem — cf. [17].

Proposition 4.1. *1. $D\mathscr{G}(u,f,\xi) \in \mathscr{L}(\mathscr{X};\mathscr{X})$ for all $(u,f,\xi) \in \mathscr{X}$.*

2. *$\mathscr{G}$ is twice continuously differentiable and $D^2\mathscr{G}$ is bounded on all bounded sets of $\mathscr{X}$.*

3. *For any $(\tilde{r},\tilde{\tau}) \in \mathscr{Y}$, $\|(\mathscr{T} - \mathscr{T}^{hp})(\tilde{r},\tilde{\tau})\|_{\mathscr{X}} \to 0$ as $h,p \to 0$.*

4. *$\|\mathscr{T} - \mathscr{T}^{hp}\|_{\mathscr{L}(\mathscr{Z},\mathscr{X})} \to 0$ as $h,p \to 0$.*

5. *A solution of* (4.19) *is regular.*

Since the items in Proposition 4.1 cover all of the assumptions of Theorem 4.2, we now have results as follows.

Theorem 4.3. *Assume that $U \in \mathscr{H}_0^1(D)$. Let $(u,\xi) \in \mathscr{H}_0^1(D) \times \mathscr{H}_0^1(D)$ be the solution of the optimality system* (4.9). *Let $(u^{hp},\xi^{hp}) \in V^{hp} \times V^{hp}$ be the solution of the discrete optimality system* (4.23) *and* (4.24). *Then*

$$\|u - u^{hp}\|_{\mathscr{H}_0^1(D)} + \|\xi - \xi^{hp}\|_{\mathscr{H}_0^1(D)} \to 0 \text{ as } h,p \to 0 .$$

Moreover, there exists $C > 0$ such that

$$\begin{aligned} &\|u - u^{hp}\|_{\mathscr{H}_0^1(D)} + \|\xi - \xi^{hp}\|_{\mathscr{H}_0^1(D)} \\ \le &Ch(\|f\|_{\mathscr{L}^2(D)} + \|u - U\|_{\mathscr{L}^2(D)}) + \sqrt{\|\rho/\hat{\rho}\|_{L^\infty(\Gamma)}} C(r_{min},N)\eta^{-\mu}, \end{aligned} \tag{4.25}$$

where the isotropic Smolyak formula based on $\hat{\rho}$-Gaussian abscissas is used.

Remark 4.1. The factor Ch in the first term on the right-hand side of (4.25) can be changed to Ch^s, $h > 1$, where s is determined by the smoothness of (u,ξ) and the degree of the approximating finite element subspace.

Similar results for the isotropic Smolyak formula based on Clenshaw-Curtis abscissas can readily be derived.

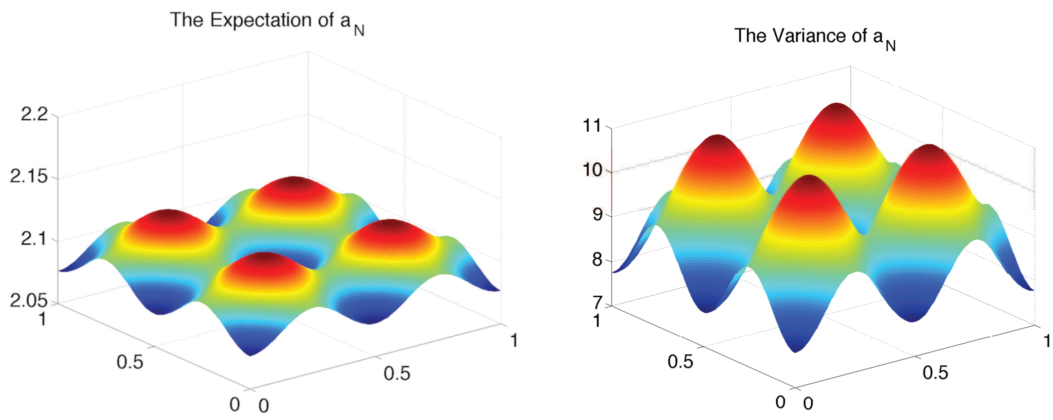

Figure 2: The expectation and variance of $a_N(x,z,\omega)$.

5. Numerical Computation of Stochastic Control Problems

This section illustrates the convergence of the sparse collocation method for a stochastic elliptic problem in two dimensions, including that the computational results are in accord with the convergence rate predicted by the theory. The constrained minimisation problem is given by

$$\mathscr{J}(u,f)=\mathbb{E}\left(\frac{1}{2}\int_D |u-U|^2\,dx+\frac{\delta}{2}\int_D |f|^2\,dx\right) \tag{5.1}$$

subject to (1.2) with $D=(0,1)^2$ and $\delta=10^{-7}$, and the reduced optimality system (4.9) is solved to get optimal solutions. We consider a deterministic desired state $U=\sin(\pi x)\sin(\pi z)$ and construct the random diffusion coefficient $a_N(x,z,\omega)$ with two-dimensional spatial dependence as

$$\begin{aligned} a_N(x,z,\omega)=&a_{min}+\exp\Big\{\left[Y_1(\omega)\cos(\pi x)+Y_3(\omega)\sin(\pi z)\right]e^{-\frac{1}{8}}\\ &+\left[Y_2(\omega)\cos(\pi x)+Y_4(\omega)\sin(\pi z)\right]e^{-\frac{1}{8}}\Big\}, \end{aligned} \tag{5.2}$$

where $a_{min}=1/100$, and the real random variables Y_i, $i=1,\cdots,4$ are independent, and have zero mean and unit variance — i.e. $\mathbb{E}[Y_i]=0$ and $\mathbb{E}[Y_iY_j]=\delta_{ij}$ for $i,j\in\mathbb{N}_+$.

The random variables Y_i, $i=1,\cdots,4$ are uniformly distributed in the interval $[-\sqrt{3},\sqrt{3}]$, and we obtain the joint probability density function ρ of $(Y_1,\cdots,Y_4)$ in this case is $(2\sqrt{3})^{-4}$ — and in our collocation method we may use either the Clenshaw-Curtis or Gauss-Legendre abscissass. For the physical domain, we use quadratic triangle finite elements.

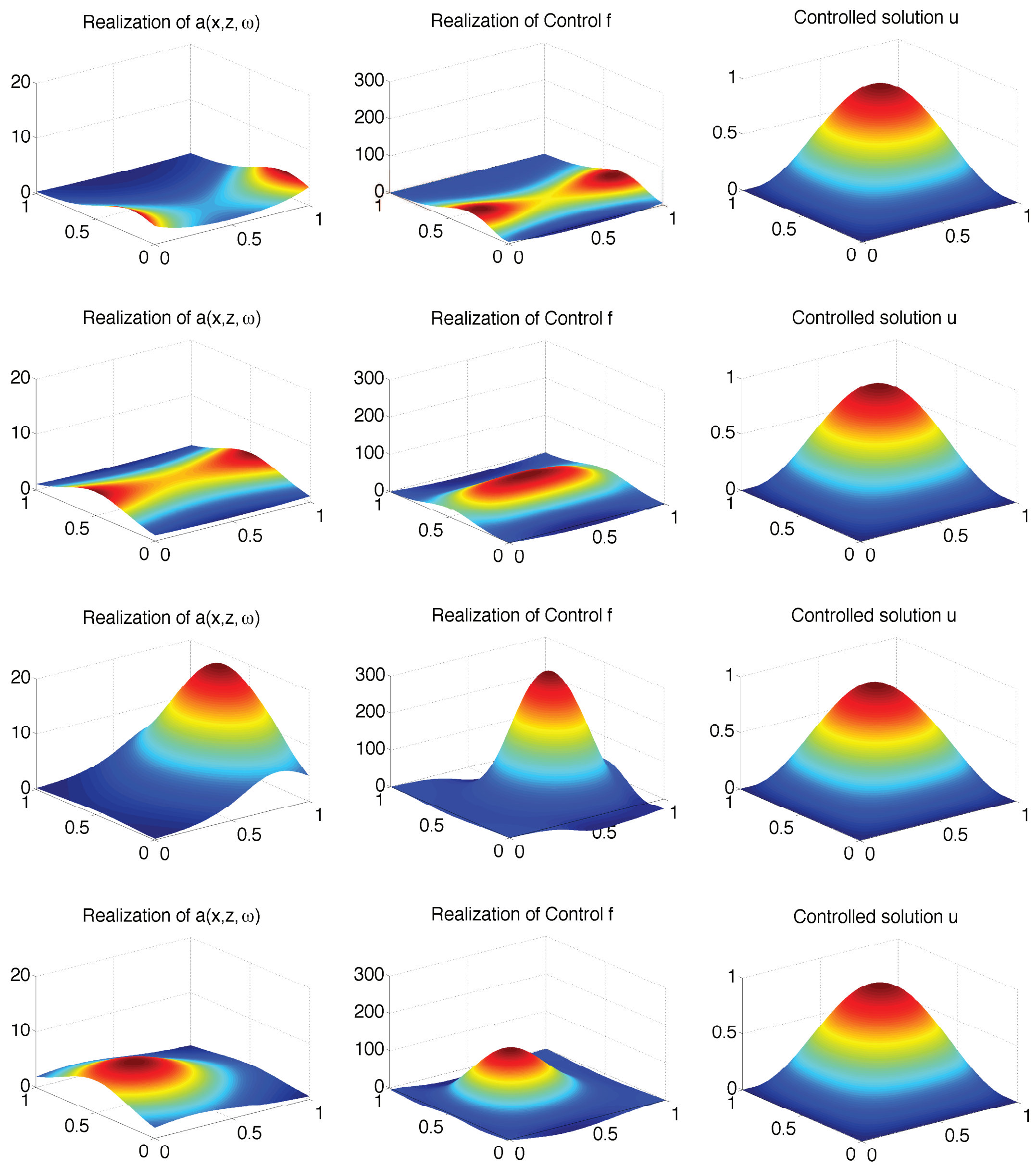

Figure 3: Some realizations of $a_N(x,z,\omega)$ (left) with corresponding distributed control $f(x,z)$ (middle) and controlled solution $u(x,z,\omega)$ (right).

5.1. Numerical results

Fig. 3 shows a realisation of the diffusivity coefficient $a(x,z,\omega)$, with corresponding distributed control f^{hp} and controlled solution $u(x,z,\omega)$. The distributed control f^{hp} changes as the controlling u^{hp} affected the random diffusivity coefficient. As expected,

Table 1: Results for Clenshaw-Curtis abscissas with dimension $N = 4$ and level 6, comparing values corresponding to changing h.

$1/h$	$\|\|u-u^{hp}\|\|_{\mathscr{H}_0^1(D)}$	Rate of conv.	$\|\|\xi-\xi^{hp}\|\|_{\mathscr{H}_0^1(D)}$	Rate of conv.
2	5.183913772882e-01		6.262470237008e-06	
4	2.553241614867e-01	2.030	3.367995092724e-06	1.859
8	1.262056585611e-01	2.023	1.277335662966e-06	2.637
16	6.290141199107e-02	2.006	5.259755611205e-07	2.429

Table 2: Number of points for Clenshaw-Curtis and Gauss-Legendre abscissas with dimension 4.

Level	level 0	level 1	level 2	level 3	level 4	level 5	level 6
# of points for CC	1	9	41	137	401	1105	7537
# of points for GL	1	9	57	289	1268	4994	

Table 3: Relative errors for optimal solutions using Clenshaw-Curtis abscissas.

Level	Relative Error for u	Relative Error for ξ	Relative Error for f
0	2.766738358942e-03	4.783631771340e-01	5.203362074889e-01
1	1.731924111923e-03	9.981821252278e-02	8.884437856187e-02
2	6.159036626895e-04	8.327814181055e-02	1.887166685990e-02
3	4.755702436742e-04	2.834413308018e-02	6.628139678764e-03
4	2.698576676732e-04	1.603426219649e-02	3.267251655201e-03
5	6.924047563378e-05	1.072228515414e-02	2.006748962780e-03

Table 4: Relative errors for optimal solutions using Gauss-Legendre abscissas.

Level	Relative Error for u	Relative Error for ξ	Relative Error for f
0	2.770736849938e-03	4.785091045750e-01	5.203267896152e-01
1	1.707906304027e-03	1.150198312016e-01	1.156558795547e-01
2	6.163433564976e-04	7.795706690576e-02	1.944239192774e-02
3	2.789940751303e-04	1.045149179460e-02	2.890386272546e-03
4	2.336437664557e-05	2.363674580206e-03	3.893077356697e-04

the controlled solution u^{hp} is similar to the desired state $U = \sin(\pi x)\sin(\pi z)$, so the expected value of u^{hp}, $\mathbb{E}[u^{hp}]$ has shape like U and the variance $\mathbb{V}ar[u^{hp}]$ is very small (about 10^{-5}). On the other hand, the variance of the control f^{hp}, $\mathbb{V}ar[f^{hp}]$ has very large values (about 10^3).

We now discuss graphs of the discrete optimal solutions that we obtained, together with the target solution and tables of relative errors in the optimal solutions, for various level Clenshaw-Curtis (CC) and Gauss-Legendre (GL) abscissas.

As shown in the Table 1, the $\mathscr{H}_0^1(D)$-errors of $||u-u^{hp}||$ and $||\xi-\xi^{hp}||$ go to zero as $h \to 0$ in the fixed level 6 and dim $N = 4$ with Clenshaw-Curtis abscissas. The rate of

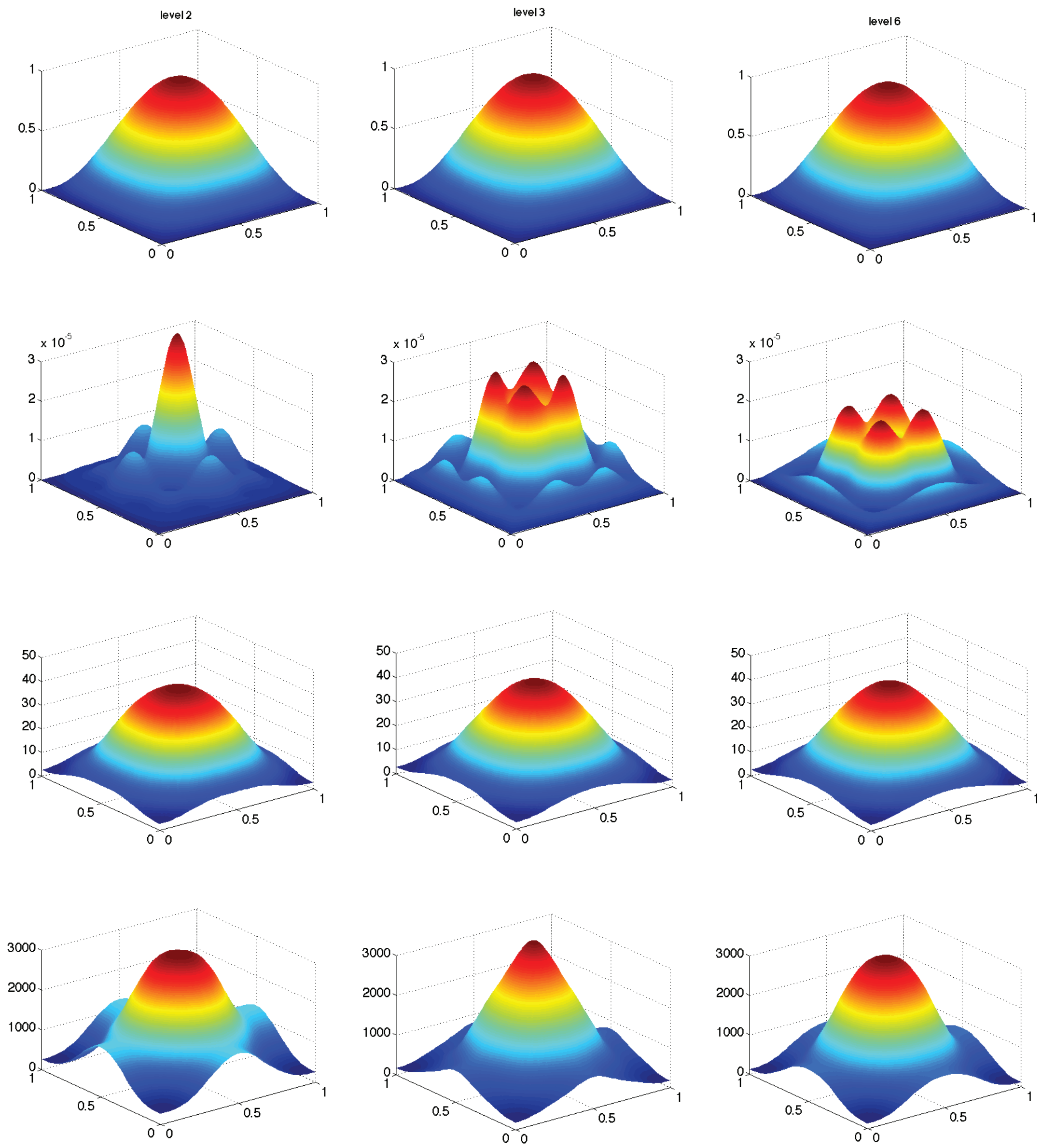

Figure 4: $\mathbb{E}[u]$ (first row), $Var[u]$ (second row), $\mathbb{E}[f]$ (third row) and $Var[f]$ (fourth row) for level 2 (first column), level 3 (second column) and level 6 (third column).

convergence becomes h^2 as h decreases in agreement with Theorem 4.3, using quadratic finite elements for the physical domain $D = (0,1)^2$.

The number of points for each method are shown in Table 2. The number of points for Clenshaw-Curtis abscissas is relatively small compared to using Gauss-Legendre abscissas at the same level. For comparable values calculated using a similar number of points, the

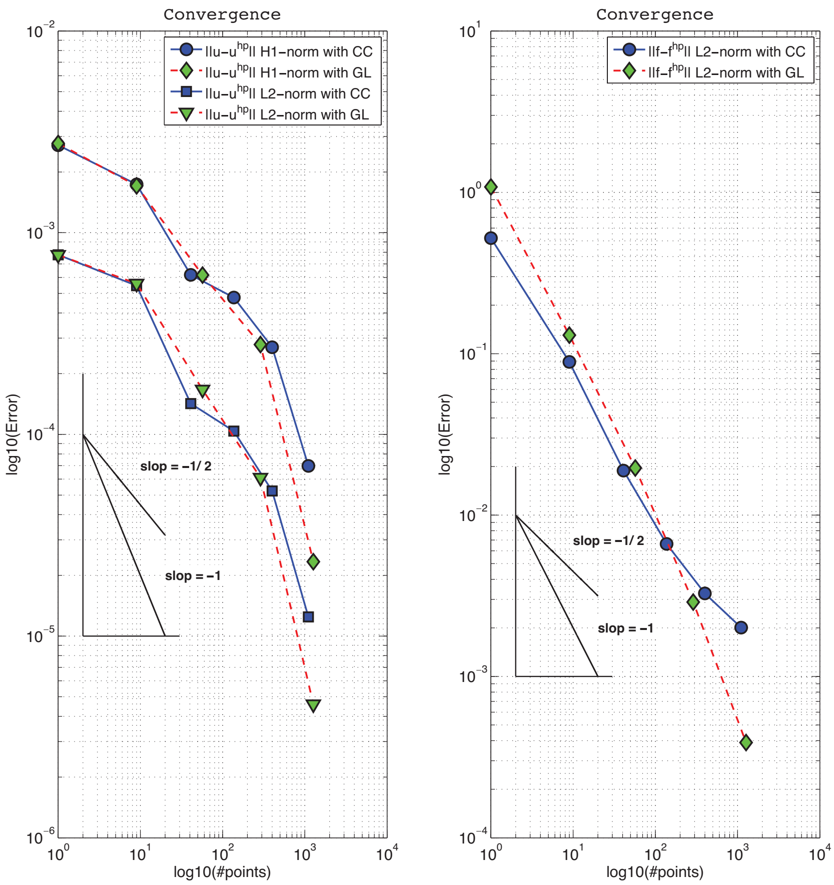

Figure 5: Convergence of the solution u^{hp} and the control f^{hp}.

level for the Clenshaw-Curtis is generally higher than for the Gauss-Legendre abscissas. To see the error corresponding to a level, we fix the step size to be $h = 1/16$ on the spatial domain and the dimension $N = 4$ on the stochastic domain, and investigate the behaviour when the level w in the Smolyak formula is increased linearly. In particular, we focus on the convergence of the discrete optimal solutions u^{hp}, ξ^{hp} and f^{hp} with respect to the level, in terms of the relative error norms. For example, for the state solution u and the Lagrange multiplier ξ we use the H^1 norm, and for our control f we use the L^2 norm. To estimate the computational relative error in the level w we approximate

$$||\mathbb{E}[\epsilon(u^{up})]|| \approx \frac{||\mathbb{E}[\mathscr{A}(w,N)\pi_h u_N - \mathscr{A}(\tilde{w},N)\pi_h u_N]||}{||\mathbb{E}[\mathscr{A}(\tilde{w},N)\pi_h u_N]||} \tag{5.3}$$

and $||\mathbb{E}[\epsilon(f^{hp})]||$ in the same way, where the $\tilde{w}$'s are 6 and 5 for the Clenshaw-Curtis and Gauss-Legendre, respectively.

We plot the convergence of the relative errors of the approximate solution $||u - u^{hp}||$ and $||f - f^{hp}||$ with respect to the number of the collocation points in Fig. 5. We ran our programs by increasing the level w on fixed spatial and stochastic domains as mentioned above, with Clenshaw-Curtis and Gauss-Legendre abscissas as shown in Fig. 5. As shown in Fig. 5, based on the values in tables 3 and 4 our results reveal that the error decreases sub-exponentially as the level w increases, confirming the theoretical convergence rates of the optimal solutions.

6. Concluding Remarks

We have successfully analysed and tested a sparse grid collocation method for an optimal control problem involving a stochastic partial differential equation and random inputs. Many mathematicians and engineers have worked on directly relevant problems for at least two decades, but the corresponding theory and analysis is not trivial. We intend to extend this work to sparse collocation calculations for the Navier-Stokes and Boussinesq fluid mechanics equations with random inputs.

Acknowledgments

The research leading to this article was supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF) (Grant Numbers NRF-2010-0010013 and NRF-2010-0026032).

References

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

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

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

[4] I. Babuska, F. Nobile and R. Temone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal. **45**, 1005-1034 (2007).

[5] I. Babuska, F. Nobile and R. Temone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Review **52**, 317-355 (2010).

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

[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**, 1251-1294 (2005).

[8] V. Barthelmann, E. Novak and K. Ritter, *High dimensional polynomial interpolation on sparse grids*, Adv. Comput. Math. **12**, 273-288 (2000).

[9] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods,* Second Edition. Springer, 2002.

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

[11] M. Crouzeix and J. Rappaz, *On Numerical Approximation in Bifurcation Theory*. Masson, Paris,1990.

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

[13] A.!Debussche, M. Fuhman and G. Tessitore, *Optimal control of a stochastic heat equation with boundary-noise and boundary-control*, ESAIM:COCV **13**, 178-205 (2007).

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

[15] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag, 1991.

[16] V. Girault and P. Raviart, *Finite Element Methods for Navier-Stokes Equations*. Springer, Berlin, 1986.

[17] M.D. Gunzburger, H.-C. Lee and J. Lee, *Error estimates of stochastic optimal Neumann boundary control problems.*, SIAM J. Numer. Anal. **49**, 1532-1552 (2011).

[18] J.L. Lions, *Optimal Control of Systems governed by Partial Differential Equations*. Springer, New York, 1971.

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

[20] F. Nobile, R. Tempone and C.G. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal. **46**, 2309-2345 (2008).

[21] F. Nobile, R. Tempone and C.G. Webster, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal. **46**, 2411-2442 (2008).

[22] C. Schwab and R.A. Todor, *Sparse finite elements for elliptic problems with stochastic loading*, Numer. Math. **95**, 707–734 (2003).

[23] S.A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR **4**, 240-243 (1963).

[24] D. Xiu and J.S. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput. **27**, 1118-1139 (2005).

[25] D. Xiu, Fast numerical methods for stochastic computations: A review, Commun. Comput. Phys. **5**, 242-272 (2009).