MATHICSE
Mathematics Institute of Computational Science and Engineering
School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report
Nr. 45.2013
December 2013

Analytic regularity and collocation approximation for PDEs with random domain deformations

J. E. Castrillon-Candas, F. Nobile, R. F. Tempone

http://mathicse.epfl.ch

Phone: +41 21 69 37648

Address:
EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland

Fax: +41 21 69 32545

ANALYTIC REGULARITY AND COLLOCATION APPROXIMATION FOR PDES WITH RANDOM DOMAIN DEFORMATIONS

JULIO E. CASTRILLON-CANDAS, FABIO NOBILE, AND RAUL F. TEMPONE

Abstract. In this work we consider the problem of approximating the statistics of a given Quantity of Interest (QoI) that depends on the solution of a linear elliptic PDE defined over a random domain parameterized by N random variables. The elliptic problem is remapped on to a corresponding PDE with a fixed deterministic domain. We show that the solution can be analytically extended to a well defined region in $\mathbb{C}^N$ with respect to the random variables. A sparse grid stochastic collocation method is then used to compute the mean and standard deviation of the QoI. Finally, convergence rates for the mean and variance of the QoI are derived and compared to those obtained in numerical experiments.

1. Introduction

In many physical processes the practicing engineer or scientist encounters the problem of optimal design under uncertainty of the underlying domain. For example, in graphine sheet nano fabrication the exact geometries of the designed patterns (e.g. nano pores) are not easy to control due to uncertainties. If there is no quantitative understanding in the involved domain uncertainty such a design may be carried out by trial and error. However, in order to accelerate the design cycle, it is essential to quantify the influence of this uncertainty on Quantities of Interest, for example the sheet stress of the graphene sheet. Other examples include lithographic process introduced in semi-conductor design [1].

Collocation and perturbation approaches have been suggested in the past as an approach to quantify the statistics of the QoI with random domains [2, 3, 1, 4, 5]. The collocation approaches proposed in [2, 3, 4] work well for large amplitude domain perturbations although suffer from the curse of dimensionality. Moreover, these works lack error estimates of the QoI with respect to the number of sparse grid points. On the other hand, the perturbations approaches introduced in [5, 1] are efficient for small domains perturbations.

In this paper we give a rigorous convergence analysis of the collocation approach based on isotropic Smolyak grids. This consists of an analysis of the regularity of the solution with respect to the parameters describing the domain perturbation. In this respect we show that the solution can be analytically extended to a well defined region in $\mathbb{C}^N$ with respect to the random variables. Moreover, we derive error estimates both in the "energy norm" as well as on functionals of the solution (Quantity of Interest) for Clenshaw Curtis abscissas that can be easily generalized to a larger class of sparse grids.

The outline of the paper is the following: In Section 2 we set up the mathematical problem and reformulate the random domain elliptic PDE problem onto a deterministic domain with random matrix coefficients. We assume that the random boundary is parameterized by N random variables.

Key words and phrases. Uncertainty Quantification, Stochastic Collocation, Stochastic PDEs, Finite Elements, Complex Analysis, Smolyak Sparse Grids.

In Section 3 we show that the solution can be analytically extended into a well defined region in $\mathbb{C}^N$. Theorem 1 is the main result of this paper. In Section 4 we setup the stochastic collocation problem and summarize several known sparse grids approaches that are used to approximate the mean and variance of the QoI. In Section 5 we assume that the random domain is truncated to $N_s \leqslant N$ random variables. We derive error estimates for the mean and variance of the QoI with respect to the finite element, sparse grid and truncation approximations. Finally, in section 7 numerical examples are presented.

2. Setup and problem formulation

Let Ω be the set of outcomes from the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is a sigma algebra of events and $\mathbb{P}$ is a probability measure. Define $L^q_P(\Omega)$, $q \in [1, \infty]$, as the space of random variables such that

$$L^q_P(\Omega) := \{v \mid \int_\Omega |v(\omega)|^q \, d\mathbb{P}(\omega) < \infty\} \text{ and } L^\infty_P(\Omega) := \{v \mid \operatorname*{ess\,sup}_{\omega \in \Omega} |v(\omega)| < \infty\},$$

where $v : \Omega \to \mathbb{R}$ be a measurable random variable.

Suppose $D(\omega) \subset \mathbb{R}^d$ is an open bounded domain with Lipschitz boundary $\partial D(\omega)$ parameterized with respect to a stochastic parameter $\omega \in \Omega$. The strong form of the problem we consider in this work is: Given $f(\cdot, \omega), a(\cdot, \omega) \in C^1(D(\omega))$ (these assumptions will be relaxed for the weak form), find $u : D(\omega) \to \mathbb{R}$ such that almost surely

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

Now, assume the diffusion coefficient satisfies the following assumption.

Assumption 1. *There exist constants a_{min} and a_{max} such that*

$$0 < a_{min} \leqslant a(x,\omega) \leqslant a_{max} < \infty \text{ for a.e. } x \in D(\omega),\ \omega \in \Omega.$$

where

$$a_{min} := \operatorname*{ess\,inf}_{x \in D(\omega), \omega \in \Omega} a(x,\omega) \quad \text{and} \quad a_{max} := \operatorname*{ess\,sup}_{x \in D(\omega), \omega \in \Omega} a(x,\omega)$$

We now state the weak formulation as:

Problem 1. *Find $u(\cdot, \omega) \in H^1_0(D(\omega))$ s.t.*

$$\int_{D(\omega)} a(x,\omega)\nabla u(x,\omega) \cdot \nabla v(x) \, dx = \int_{D(\omega)} f(x,\omega) v(x) \, dx \quad \forall v \in H^1_0(D(\omega)) \text{ a.s. in } \Omega, \tag{1}$$

where $f(\cdot, \omega) \in L^2(D(\omega))$ for a.e. $\omega \in \Omega$.

Under Assumption 1 the weak formulation has a unique solution up to a zero-measure set in Ω.

2.1. **Reformulation onto a fixed Domain.** Now, assume that given any $\omega \in \Omega$ the domain $D(\omega)$ can be mapped to a reference domain $U \subset \mathbb{R}^d$ with Lipschitz boundary through a random map $F : U \times \Omega \to \mathbb{R}^d$, where we assume that F is one-to-one and the determinant of the Jacobian $|\partial F(\cdot, \omega)| \in W^{1,\infty}(U)$ for any $\omega \in \Omega$. Furthermore we assume that $|\partial F|$ is uniformly greater than zero almost surely. We will, however, make the following equivalent assumption.

Assumption 2. *Given a one-to-one map $F : U \times \Omega \to \mathbb{R}^d$ there exist constants $\mathbb{F}_{min}$ and $\mathbb{F}_{max}$ such that*

$$0 < \mathbb{F}_{min} \leqslant \sigma_{min}(\partial F(\omega)) \text{ and } \sigma_{max}(\partial F(\omega)) \leqslant \mathbb{F}_{max} < \infty$$

almost everywhere in U and almost surely in Ω. We have denoted by $\sigma_{min}(\partial F(\omega))$ (and $\sigma_{max}(\partial F(\omega))$) the minimum (respectively maximum) singular value of the Jacobian $\partial F(\omega)$.

In the rest of the paper we shall drop repeating a.s. in Ω and a.e. in U unless disambiguation is needed.

Lemma 1. *Under Assumptions 2 it is immediate to prove the following results:*
i) $L^2(D(\omega))$ and $L^2(U)$ are isomorphic.
ii) $H^1(D(\omega))$ and $H^1(U)$ are isomorphic.

Problem 1 can be reformulated with respect to the fixed reference domain U. From the chain rule we have that for any $v \in C^1(D(\omega))$

$$\nabla v = \partial F^{-T} \nabla (v \circ F). \tag{2}$$

By a change of variables, the weak form can now be posed as:

Problem 2. *Find $u \circ F \in H^1_0(U)$ s.t.*

$$B(\omega; u \circ F, v) = l(\omega; v), \;\; \forall v \in H^1_0(U) \tag{3}$$

where for any $v, s \in H^1_0(U)$

$$B(\omega; s, v) := \int_U (a \circ F(\cdot, \omega) \nabla s^T C^{-1}(\cdot, \omega) \nabla v |\partial F(\cdot, \omega)|,$$
$$l(\omega; v) := \int_U (f \circ F(\cdot, \omega)) v \; |\partial F(\cdot, \omega)|,$$

$f \circ F \in L^2(U)$, and $C = \partial F^T \partial F$.

The following lemma gives the conditions under which Problem 2 is well posed.

Lemma 2. *Given that Assumptions 1 and 2 are satisfied then there exists a.s. a unique solution to Problem 2, which coincides with the solution to Problem 1, and*

$$\|\nabla u\|_{L^2(D(\omega))} \leqslant \frac{\mathbb{F}_{max}^{3d/2+2}}{a_{min} \mathbb{F}_{min}^{d+1}} \|f \circ F\|_{L^2(U)} C_P(U)$$

where $C_P(U)$ is the Poincaré constant of the reference domain U.

Proof. From Assumption 2 we have that

$$|\partial F| = \sqrt{|C|} = \sqrt{\Pi_{i=1}^d \lambda(C)} = \Pi_{i=1}^d \sigma_i(\partial F).$$

therefore $\mathbb{F}_{min}^d \leqslant |\partial F| \leqslant \mathbb{F}_{max}^d$. Furthermore, from Assumption 2 we have that

$$\lambda_{min}((a \circ F) C^{-1} |\partial F|) \geqslant a_{min} \mathbb{F}_{min}^d \lambda_{min}(C^{-1}) = a_{min} \mathbb{F}_{min}^d \mathbb{F}_{max}^{-2} > 0.$$

and

$$\lambda_{max}((a \circ F) C^{-1} |\partial F|) \leqslant a_{max} \mathbb{F}_{max}^d \lambda_{max}(C^{-1}) = a_{max} \mathbb{F}_{max}^d \mathbb{F}_{min}^{-2} < \infty.$$

Thus Problem 2 is uniformly continuous and coercive, and from the Lax-Milgram theorem there exists a.s. a unique solution. The equivalence between Problems 1 and 2 is an immediate consequence of the chain rule and the isomorphism between $H^1_0(U)$ and $H^1_0(D(\omega))$ (Lemma 1).

From the Cauchy-Schwartz inequality we obtain

$$\begin{aligned}\lambda_{min}((a\circ F)C^{-1}|\partial F|)\|\nabla(u\circ F)\|^2_{L^2(U)} &\leqslant |B(\omega;u\circ F,u\circ F)| = |l(\omega;u\circ F)| \\ &\leqslant \int_U |f\circ F||u\circ F||\partial F| \\ &\leqslant \|f\circ F\|_{L^2(U)}\|u\circ F\|_{L^2(U)}\mathbb{F}^d_{max}.\end{aligned}$$

From the Poincaré inequality ($\|u\circ F\|_{L^2(U)} \leqslant C_P(U)\|\nabla(u\circ F)\|_{L^2(U)}$) we obtain

$$\|\nabla(u\circ F)\|_{L^2(U)} \leqslant \frac{\|f\circ F\|_{L^2(U)}C_P(U)\mathbb{F}^d_{max}}{a_{min}\mathbb{F}^d_{min}\mathbb{F}^{-2}_{max}}.$$

From (2) we obtain that $\forall v \in H^1_0(D(\omega))$

$$\|\nabla v\|_{L^2(D(\omega))} \leqslant \mathbb{F}^{d/2}_{max}\mathbb{F}^{-1}_{min}\|\nabla(v\circ F)\|_{L^2(U)},$$

thus

$$\|\nabla u\|_{L^2(D(\omega))} \leqslant \frac{\mathbb{F}^{3/2d+2}_{max}}{a_{min}\mathbb{F}^{d+1}_{min}}\|f\circ F\|_{L^2(U)}C_P(U).$$

□

Remark 1. *For many practical applications the non-zero Dirichlet boundary value problem is more interesting. We can easily extend the stochastic domain problem to non-zero Dirichlet boundary conditions.*

Suppose $u\circ F \in H^1(U)$ is the weak solution to the following boundary valued problem: Find $u\circ F$ such that

$$\begin{aligned}-\nabla\cdot((a\circ F)C^{-1}|\partial F|\nabla u\circ F) &= f\circ F \text{ in } U \\ u\circ F\,|_U &= g\circ F \text{ on } \partial U\end{aligned}$$

a.s. in Ω, where $f\circ F \in L^2(U)$ and $g\circ F, a\circ F \in W^{1,\infty}(U)$.

Since U is bounded and Lipschitz there exists a bounded linear operator $T : H^{1/2}(\partial U) \to H^1(U)$ such that $\forall \hat{g} \in H^{1/2}(\partial U)$ we have that $\mathbf{w} := T\hat{g} \in H^1(U)$ satisfies $\mathbf{w}|_{\partial U} = \hat{g}$ almost surely. The weak formulation can now be posed as ([6] *chapter 6, p297):*

Problem 3. *Given that $f\circ F \in L^2(U)$ find $\tilde{u}\circ F \in H^1_0(U)$ s.t.*

$$B(\omega;\tilde{u}\circ F,v) = \tilde{l}(\omega;v),\;\; \forall v \in H^1_0(U)$$

almost surely, where $\tilde{l}(\omega;v) := \int_U f\circ F|\partial F|v - L(\mathbf{w},v)$, $L(\mathbf{w},v) := a\circ F\nabla\mathbf{w}^T C^{-1}|\partial F|\nabla v$ and $\mathbf{w} = T(g\circ F)$.

The weak solution $u\circ F \in H^1(U)$ for the non-zero Dirichlet boundary value problem is simply obtained as $u\circ F = \tilde{u}\circ F + \mathbf{w}$.

2.1.1. *Quantity of Interest and the Adjoint problem.* In practice we are interested in computing the statistics of a Quantity of Interest (QoI) over the stochastic domain or a subdomain of it. We consider QoI of the form

$$Q(u) := \int_{\bar{D}} q(x)u(x,\omega)\,dx \tag{4}$$

($g \in C^\infty(D(\omega))$) over the region $\bar{D} \subset D(\omega)$ for any $\omega \in \Omega$. Furthermore we assume that we can always construct a mapping F s.t. $\partial F|_{\bar{D}} = I$ so that $\bar{U} = F^{-1}(\bar{D})$ does not depend on the parameter $\omega \in \Omega$ i.e.

$$Q(u \circ F) = \int_{\bar{U}} (q \circ F) u \circ F \; dx = Q(u). \tag{5}$$

In this paper we restrict our attention to the computation of the mean $\mathbb{E}[Q]$ and variance $Var[Q] := \mathbb{E}[Q^2] - \mathbb{E}[Q]^2$ given that the domain deformation is parameterized by a stochastic random vector.

We first assume that $Q : H_0^1(U) \to \mathbb{R}$ is a bounded linear functional. The influence function can be computed as:

Problem 4. *Find $\varphi \in H_0^1(U)$ such that $\forall v \in H_0^1(U)$*

$$B(\omega; v, \varphi) = Q(v) \tag{6}$$

a.s. in Ω.

Now, assume that $dist(\bar{D}, \partial D) \geqslant \delta$ for some $\delta > 0$. We can now pick $\mathbf{w} = T(g \circ F)$ such that $Q(\tilde{u} \circ F) = Q(u \circ F)$, i.e $Q(\mathbf{w}) = 0$. Therefore, we have that

$$Q(u \circ F) = Q(\tilde{u} \circ F) = Q(u \circ F - Tg \circ F) = B(\omega; \tilde{u} \circ F, \varphi).$$

2.2. **Domain Parameterization.** Let $Y := [Y_1, \ldots, Y_N]$ be a N valued random vector measurable in $(\Omega, \mathcal{F}, \mathbb{P})$ taking values on $\Gamma := \Gamma_1 \times \cdots \times \Gamma_N \subset \mathbb{R}^N$ and $\mathcal{B}(\Gamma)$ be the Borel $\sigma-$algebra.

Define the induced measure μ_Y on $(\Gamma, \mathcal{B}(\Gamma))$ as $\mu_Y := \mathbb{P}(Y^{-1}(A))$ for all $A \in \mathcal{B}(\Gamma)$. Assuming that the induced measure is absolutely continuous with respect to the Lebesgue measure defined on Γ, then there exists a density function $\rho(\mathbf{y}) : \Gamma \to [0, +\infty)$ such that for any event $A \in \mathcal{B}(\Gamma)$

$$\mathbb{P}(Y \in A) := \mathbb{P}(Y^{-1}(A)) = \int_A \rho(\mathbf{y}) \, d\mathbf{y},$$

Now, for any measurable function $Y \in L_P^1(\Gamma)$ we let the expected value be defined as

$$\mathbb{E}[Y] = \int_\Gamma \mathbf{y} \rho(\mathbf{y}) \, d\mathbf{y}.$$

Remark 2. *In Sections 4 and 5 we shall use an alternative density function $\hat{\rho}(\mathbf{y}) : \Gamma \to [0, +\infty)$. To disambiguate the functional space $L_P^q(\Omega)$ with respect to $\rho(\mathbf{y})$ and $\hat{\rho}(\mathbf{y})$ we shall refer to $L_{\hat{\rho}}^q(\Omega)$ as the space $L_P^q(\Omega)$ with respect to the density function $\hat{\rho}(\mathbf{y})$.*

The mapping $F(\cdot, \omega) : U \to D(\omega)$ can be parameterized in many forms. In this paper we restrict our attention to the following class of mappings:

Suppose that $U_i \subset U \subset \mathbb{R}^d$, $i = 1, \ldots, M$, is a collection of non overlapping open elements (square, triangular, tetrahedral, nurbs, etc) in $\mathbb{R}^d$ such that $U := \cup_{i=1}^M \overline{U}_i$ forms a Lipschitz bounded domain.

For each element U_i, $i = 1, \ldots, M$ suppose we have a map $F_i : U_i \times \Omega \to \mathbb{R}^d$ that satisfies Assumption 2. Now, let $D_i(\omega) \subset \mathbb{R}^d$ be the image of $F_i(U_i, \omega)$ and denote $D(\omega) := \cup_{i=1}^M D_i(\omega)$. Assume that $D(\omega)$ is a conformal mesh.

Assumption 3. *For each open element $U_i \subset \mathbb{R}^d$, $i = 1, \ldots, M$, the map $F_i : U_i \times \Omega \to \mathbb{R}$ has the form*

$$F_i(x, \omega) := x + q_i(x, \omega),$$

where

$$q_i(x,\omega) := e_i(x,\omega)\hat{v}_i(x)$$

a.s. in Ω*, with* $\hat{v}_i : U_i \to \mathbb{R}^d$*,* $\hat{v}_i \in C^1(U_i)$*, and* $e_i(x,\omega) : U_i \times \Omega \to \mathbb{R}$*. Assume that for each* $i = 1,\dots,M$ *the maps* $F_i : U_i \times \Omega \to D_i(\omega)$ *are one-to-one almost everywhere.*

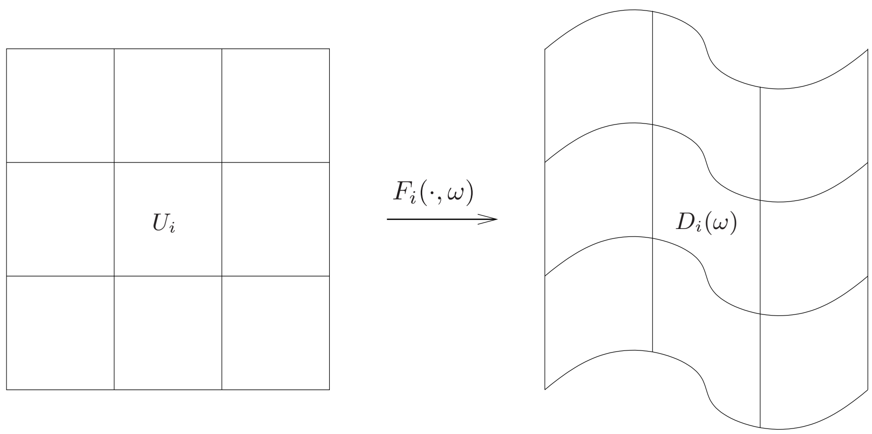


FIGURE 1. Cartoon example of stochastic domain realization from Reference elements.

The next step is to characterize the stochastic perturbation variables $e_1,\dots,e_M$. Without loss of generality we characterize only a single stochastic perturbation $e(x,\omega) : \tilde{U} \times \Omega \to D(\omega)$ for a single generic element $\tilde{U}$ with the following parameterization:

$$e(x,\omega) := \sum_{l=1}^{N} \sqrt{\lambda_l}\, b_l(x) Y_l(\omega)$$

Denote $Y := [Y_1,\dots,Y_N]$, and for $n = 1,\dots,N$, let $\Gamma_n \equiv Y_n(\Omega)$, $\mathbb{E}[Y_n] = 0$, $\mathbb{E}[Y_n^2] = 1$. Furthermore denote $\Gamma := \prod_{n=1}^{N} \Gamma_n$, and $\rho(\mathbf{y}) : \Gamma \to \mathbb{R}_+$ as the joint probability density of Y. In addition, we make the following assumptions:

Assumption 4. *(1)* $n = 1,\dots,N$, $\Gamma_n \equiv [-1,1]$
(2) $b_1,\dots,b_N \in C^\infty(\tilde{U})$
(3) $\|b_l\sqrt{\lambda_l}\|_{L^\infty(\tilde{U})}$ *are monotonically decreasing for* $l = 1, 2, \dots N$.

3. ANALYTICITY

The analytic extension of the solution of Problem 3 with respect to the images of the stochastic variables provides us a form to bound the approximation error of the collocation scheme. For notational simplicity we only derive the analyticity of the solution u with respect to the random variables $[Y_1,\dots,Y_N]$ parameterizing one simple perturbation field $F(x,\omega) : \tilde{U} \times \omega \to \mathbb{R}^d$ as $F(x,\omega) = x + e(x,\omega)\hat{v}(x)$.

In this analysis we consider only the homogeneous Dirichlet case since the extension to the non-homogeneous case is straightforward. First, we establish some notation and assumptions.

From the stochastic model formulated in Section 2 the Jacobian ∂F is written as

$$\partial F(x,\omega) = I + \sum_{l=1}^{N} B_l(x)\sqrt{\lambda_l}Y_l(\omega) \tag{7}$$

with

$$B_l(x) := b_l(x)\partial\hat{v}(x) + \begin{bmatrix} \frac{\partial b_l(x)}{\partial x_1}\hat{v}_1(x) & \frac{\partial b_l(x)}{\partial x_2}\hat{v}_1(x) & \dots & \frac{\partial b_l(x)}{\partial x_d}\hat{v}_1(x) \\ \frac{\partial b_l(x)}{\partial x_1}\hat{v}_2(x) & \frac{\partial b_l(x)}{\partial x_2}\hat{v}_2(x) & \dots & \frac{\partial b_l(x)}{\partial x_d}\hat{v}_2(x) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial b_l(x)}{\partial x_1}\hat{v}_d(x) & \frac{\partial b_l(x)}{\partial x_2}\hat{v}_d(x) & \dots & \frac{\partial b_l(x)}{\partial x_d}\hat{v}_d(x) \end{bmatrix}$$

where ∂v is the Jacobian of $v(x)$.

Assumption 5. *(a) $a \circ F$ is only a function of $x \in \tilde{U}$ and independent of $\omega \in \Omega$.*
(b) There exists $0 < \tilde{\delta} < 1$ such that $\sum_{l=1}^{N} \|B_l(x)\|_2\sqrt{\lambda_l} \leqslant 1 - \tilde{\delta}$, $\forall x \in \tilde{U}$.
(c) Assume that $f : \mathbb{R}^d \to \mathbb{R}$ and $\mathbf{w} : \mathbb{R}^d \to H^1(U)$ can be analytically extended in $\mathbb{C}^d$.

Remark 3. *Assumption 5 (a) restricts $a(x,\omega)$ to be a constant along the direction $v(x)$. This assumption simplifies the presentation of this section.*

We now extend the mapping $\partial F(\mathbf{y}) = I + R(x,\mathbf{y})$, with $R(x,\mathbf{y}) := \sum_{l=1}^{N} \sqrt{\lambda_l}B_l(x)y_l$, to the complex plane. First, for any $0 < \beta < \tilde{\delta}$ define the following region in $\mathbb{C}^N$:

$$\Theta_\beta := \left\{ \mathbf{z} \in \mathbb{C}^N ; \, \mathbf{z} = \mathbf{y} + \mathbf{w}, \, \mathbf{y} \in [-1,1]^N, \, \sum_{l=1}^{N} \sup_{x \in \tilde{U}} \|B_l(x)\|_2 \sqrt{\lambda_l}|w_l| \leqslant \beta \} \right\}. \tag{8}$$

Note that in the rest of the section for sake of simplicity we shall refer to $R(x,\mathbf{y})$ or $R(x,\mathbf{z})$ as $R(\mathbf{y})$ or $R(\mathbf{z})$ unless emphasis is needed. We shall now prove several lemmas that will be useful to prove the main results (Theorem 1).

Lemma 3. *Under Assumption 5 we have that $\forall \mathbf{y} \in [-1,1]^N$ and $x \in \tilde{U}$*

i) $\sigma_{max}(\partial F(\mathbf{y})) \leqslant 2 - \tilde{\delta}$,
ii) $\sigma_{min}(\partial F(\mathbf{y})) \geqslant \tilde{\delta}$,
iii) $(2-\tilde{\delta})^d \geqslant det(\partial F(\mathbf{y})) \geqslant \tilde{\delta}^d$.

Proof. i) $\|\partial F(\mathbf{y})\|_2 \leqslant 1 + \sup_{x\in\tilde{U}} \sum_{l=1}^{N} \|B_l(x)\|_2\sqrt{\lambda_l} \leqslant 2 - \tilde{\delta}$.
ii) $\sigma_{max}(\partial F(\mathbf{y}) - I) = \| \sum_{l=1}^{N} B_l(x)\sqrt{\lambda_l}y_l\|_2 \leqslant 1 - \tilde{\delta} \Rightarrow \sigma_{min}(\partial F(\mathbf{y})) = \sigma_{min}(I + \partial F(\mathbf{y}) - I) \geqslant \sigma_{min}(I) - \sigma_{max}(\partial F(\mathbf{y}) - I) \geqslant 1 - (1 - \tilde{\delta}) = \tilde{\delta}$
iii) The result follows from the following fact: If $A \in \mathbb{C}^{d\times d}$ we have that $\sigma_{min}(A) \leqslant |\lambda_l(A)| \leqslant \sigma_{max}(A)$ for all $l = 1,\dots,N$.

□

Lemma 4. *Let $0 < \beta < \tilde{\delta}\frac{\log 2}{d+\log 2}$ and $\alpha = 2 - exp(\frac{d\beta}{\tilde{\delta}-\beta}) > 0$ then $\forall \mathbf{z} \in \Theta_\beta$ and $\forall x \in \tilde{U}$ we have that $det(\partial F(\mathbf{z}))$ is analytic and*

i) $|det(\partial F(\mathbf{z}))| \geqslant \tilde{\delta}^d\alpha$,
ii) $|det(\partial F(\mathbf{z}))| \leqslant (2-\tilde{\delta})^d(2-\alpha)$,
iii) $\mathrm{Re}\, det(\partial F(\mathbf{z})) \geqslant \tilde{\delta}^d\alpha$, $|\mathrm{Im}\, det(\partial F(\mathbf{z}))| \leqslant (2-\tilde{\delta})^d(1-\alpha)$.

Proof. For all $\mathbf{z} \in \Theta_\beta$ we have that

$$\partial F(x,\mathbf{z}) = I + \sum_{l=1}^{N} B_l(x)\sqrt{\lambda_l} y_l + \sum_{l=1}^{N} B_l(x)\sqrt{\lambda_l} w_l = I + R(\mathbf{y}) + R(\mathbf{w})$$

and let $Q(\mathbf{y},\mathbf{w}) = I + \partial F(\mathbf{y})^{-1}R(\mathbf{w})$ so that $\partial F(\mathbf{z}) = \partial F(\mathbf{y})Q(\mathbf{y},\mathbf{w})$.

We now study $det(Q(\mathbf{y},\mathbf{w}))$ for all $\mathbf{z} \in \Theta_\beta$ by using the following identity [7]: If $A \in \mathbb{C}^{d\times d}$ and $\sigma_{max}(A) < 1$ then

$$det(I+A) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!}\left(-\sum_{j=1}^{\infty}\frac{(-1)^j}{j} tr(A^j)\right)^k$$

It follows that

$$\begin{aligned}
|det(Q(\mathbf{y},\mathbf{w})) - 1| &\leqslant \sum_{k=1}^{\infty}\frac{1}{k!}\left(\sum_{j=1}^{\infty}\frac{1}{j}|tr((\partial F(\mathbf{y})^{-1}R(\mathbf{w}))^j)|\right)^k \\
&\leqslant \sum_{k=1}^{\infty}\frac{1}{k!}\left(\sum_{j=1}^{\infty}\frac{1}{j}\sum_{l=1}^{d}|\lambda_l(\partial F(\mathbf{y})^{-1}R(\mathbf{w}))|^j\right)^k \\
&\leqslant \sum_{k=1}^{\infty}\frac{1}{k!}\left(\sum_{j=1}^{\infty}\frac{1}{j} d\sigma_{min}^{-j}(\partial F(\mathbf{y}))\sigma_{max}^{j}(R(\mathbf{w}))\right)^k \\
&\leqslant \sum_{k=1}^{\infty}\frac{1}{k!}\left(\sum_{j=1}^{\infty} d(\beta/\tilde{\delta})^j\right)^k \\
&\text{(From Lemma 3)} \\
&= \sum_{k=1}^{\infty}\frac{1}{k!}\left(\frac{d(\beta/\tilde{\delta})}{1-\beta/\tilde{\delta}}\right)^k = \exp\left(\frac{d\beta}{\tilde{\delta}-\beta}\right) - 1.
\end{aligned}$$

and $|det(Q(\mathbf{y},\mathbf{w}))| \leqslant \exp\left(\frac{d\beta}{\tilde{\delta}-\beta}\right)$.

Now, it follows that if $\beta < \frac{\tilde{\delta}\log 2}{d+\log 2}$ then

i) $|det(Q(\mathbf{y},\mathbf{w}))| \geqslant 1 - |det(Q(\mathbf{y},\mathbf{w})) - 1| \geqslant 2 - \exp\left(\frac{d\beta}{\tilde{\delta}-\beta}\right) = \alpha > 0$,
ii) $\operatorname{Re} det(Q(\mathbf{y},\mathbf{w})) \geqslant 1 - |det(Q(\mathbf{y},\mathbf{w})) - 1| \geqslant \alpha > 0$,
iii) $|\operatorname{Im} det(Q(\mathbf{y},\mathbf{w}))| \leqslant |det(Q(\mathbf{y},\mathbf{w})) - 1| \leqslant 1 - \alpha$.

Finally we have that $det(\partial F(\mathbf{z})) = det(\partial F(\mathbf{y}))det(Q(\mathbf{y},\mathbf{w}))$. It is easy now to see that $det(\partial F(\mathbf{z}))$ is analytic $\forall \mathbf{z} \in \Theta_\beta$ since $det(Q(\mathbf{y},\mathbf{w}))$ is a finite polynomial of $\mathbf{w}$. The rest of the result follows by applying Lemma 3. □

Lemma 5. *Let* $G(\mathbf{z}) := (a \circ F)det(\partial F(\mathbf{z}))\partial F^{-1}(\mathbf{z})\partial F^{-T}(\mathbf{z})$ *and suppose*

$$0 < \beta < \min\{\tilde{\delta}\frac{\log(2-\gamma)}{d+\log(2-\gamma)}, \sqrt{1+\tilde{\delta}^2/2} - 1\}$$

where $\gamma := \frac{(2-\tilde{\delta})^d}{\tilde{\delta}^d + (2-\tilde{\delta})^d}$ *then* $\operatorname{Re} G(\mathbf{z})$ *is positive definite* $\forall \mathbf{z} \in \Theta_\beta$ *and* *(a)* $\lambda_{min}(\operatorname{Re} G(\mathbf{z})^{-1}) \geqslant B(\tilde{\delta}, \beta, d, a_{max}) > 0$ *where*

$$B(\tilde{\delta}, \beta, d, a_{max}) := \frac{\tilde{\delta}^d((\tilde{\delta} - \beta)^2 - \beta^2) + 2(\beta(2 + (\beta - \tilde{\delta}))}{a_{max}(2 - \tilde{\delta})^{2d}(2 - \alpha)^2}.$$

(b) $\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1}) \leqslant D(\tilde{\delta}, \beta, d, a_{min}) < \infty$ *where*

$$D(\tilde{\delta}, \beta, d, a_{min}) := \frac{1}{a_{min}\tilde{\delta}^{2d}\alpha^2}\Big[(2 - \tilde{\delta})^d(2 - \alpha)(\tilde{\delta} + \beta)^2 \\ +2(1 - \tilde{\delta})^d(2 - \alpha)(\beta(2 + (\beta - \tilde{\delta}))\Big].$$

(c) $\sigma_{max}(\operatorname{Im} G(\mathbf{z})^{-1}) \leqslant C(\tilde{\delta}, \beta, d, a_{min}) < \infty$ *where*

$$C(\tilde{\delta}, \beta, d, a_{min}) := \frac{1}{a_{min}\tilde{\delta}^{2d}\alpha^2}\Big[(2 - \tilde{\delta})^d(2 - \alpha)2\beta(2 + (\beta - \tilde{\delta})) \\ +(2 - \tilde{\delta})^d(1 - \alpha)((2 - \tilde{\delta}) + \beta)^2 + \beta^2)\Big].$$

Proof. **(a)** To simplify the proof we use the property that if $\operatorname{Re} G^{-1}(\mathbf{z})$ is positive definite then $\operatorname{Re} G(\mathbf{z})$ is positive definite (From (b) in [8]) , but first we derive bounds for $\operatorname{Re} \partial F(\mathbf{z})^T \partial F(\mathbf{z})$ and $\operatorname{Im} \partial F(\mathbf{z})^T \partial F(\mathbf{z})$. For all $\mathbf{z} \in \Theta_\beta$ we have that

$$\begin{aligned}\operatorname{Re} \partial F(\mathbf{z})^T \partial F(\mathbf{z}) &= \operatorname{Re}[(I + R(\mathbf{y}) + R(\mathbf{w}))^T (I + R(\mathbf{z}) + R(\mathbf{w}))] \\ &= (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T (I + R(\mathbf{y}) + R_r(\mathbf{w})) - R_i(\mathbf{w})^T R_i(\mathbf{w}).\end{aligned}$$

where $R(\mathbf{w}) = R_r(\mathbf{w}) + iR_i(\mathbf{w})$. By applying the dual Lidskii inequality (if $A, B \in \mathbb{C}^{d\times d}$ are Hermitian then $\lambda_{min}(A + B) \geqslant \lambda_{min}(A) + \lambda_{min}(B)$) we obtain

$$\begin{aligned}\lambda_{min}(\operatorname{Re} \partial F(\mathbf{z})^T \partial F(\mathbf{z})) &\geqslant \lambda_{min}((I + R(\mathbf{y}) + R_r(\mathbf{w}))^T (I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\ &\quad - \lambda_{max}(R_i(\mathbf{w})^T R_i(\mathbf{w})) \\ &= \sigma^2_{min}(I + R(\mathbf{y}) + R_r(\mathbf{w})) - \sigma^2_{max}(R_i(\mathbf{w})) \\ &\geqslant (\sigma_{min}(I + R(\mathbf{y})) - \sigma_{max}(R_r(\mathbf{w})))^2 - \sigma^2_{max}(R_i(\mathbf{w})) \\ &\geqslant (\tilde{\delta} - \beta)^2 - \beta^2.\end{aligned} \tag{9}$$

It follows that if $\beta < \tilde{\delta}/2$ then

$$\lambda_{min}(\operatorname{Re} \partial F(\mathbf{z})^T \partial F(\mathbf{z})) \geqslant \tilde{\delta}(\tilde{\delta} - 2\beta) > 0.$$

and is positive definite. We see that for all $\mathbf{z} \in \Theta_\beta$,

$$\begin{aligned}|\lambda_{min}(\operatorname{Im} \partial F(\mathbf{z})^T \partial F(\mathbf{z}))| &\leqslant |\lambda_{max}(R_i(\mathbf{w})^T (I + R(\mathbf{y}) + R_r(\mathbf{w})) \\ &\quad + (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w}))| \\ &\leqslant \sigma_{max}(R_i(\mathbf{w})^T (I + R(\mathbf{y}) + R_r(\mathbf{w})) \\ &\quad + (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w})) \\ &\leqslant 2\sigma_{max}(R_i(\mathbf{w}))\sigma_{max}(I + R(\mathbf{y}) + R_r(\mathbf{w})) \\ &\leqslant 2\beta(2 + (\beta - \tilde{\delta})).\end{aligned} \tag{10}$$

We now have that

$$
\begin{aligned}
\lambda_{min}(a^{-1}\operatorname{Re}(\xi^{-1}(\mathbf{z})\partial F(\mathbf{z})^T\partial F(\mathbf{z}))) &\geqslant \frac{1}{a_{max}|\xi(\mathbf{z})|^2}\lambda_{min}(\xi_R(\mathbf{z})\operatorname{Re}F(\mathbf{z})^T\partial F(\mathbf{z}) \\
&\quad + \xi_I(\mathbf{z})\operatorname{Im}F(\mathbf{z})^T\partial F(\mathbf{z}))) \\
&\geqslant \frac{1}{a_{max}|\xi(\mathbf{z})|^2}(\xi_R(\mathbf{z})\lambda_{min}(\operatorname{Re}\partial F(\mathbf{z})^T\partial F(\mathbf{z})) \\
&\quad - |\xi_I(\mathbf{z})||\lambda_{min}(\operatorname{Im}\partial F(\mathbf{z})^T\partial F(\mathbf{z}))|)),
\end{aligned} \tag{11}
$$

where $\xi(\mathbf{z}) := \xi_R(\mathbf{z}) + i\xi_I(\mathbf{z}) = det(I + R(\mathbf{z}))$. From Lemma 4 we have that $|\xi(\mathbf{z})|^{-1} \geqslant (2-\tilde{\delta})^{-d}(2-\alpha) > 0$ whenever $\mathbf{z} \in \Theta_\beta$.

From Lemma 3 (iii) if $\beta < \frac{\tilde{\delta}\log(2-\gamma)}{d+\log(2-\gamma)}$, $\gamma := \frac{(2-\tilde{\delta})^d + 2\tilde{\delta}^d}{\tilde{\delta}^d + (2-\tilde{\delta})^d}$, then $\xi_R(\mathbf{z}) > |\xi_I(\mathbf{z})|$, $\forall \mathbf{z} \in \Theta_\beta$. From inequalities (9) and (10) we have that if $\beta < \sqrt{1 + \tilde{\delta}^2/2} - 1$ then $\lambda_{min}(\operatorname{Re}\partial F(\mathbf{z})^T\partial F(\mathbf{z})) > |\lambda_{min}(\operatorname{Im}\partial F(\mathbf{z})^T\partial F(\mathbf{z}))|$ and $\lambda_{min}(\operatorname{Re}G(\mathbf{z})^{-1}) \geqslant B(\tilde{\delta}, \beta, d, a_{max}) > 0$ where

$$
B(\tilde{\delta}, \beta, d, a_{max}) := \frac{\tilde{\delta}^d\alpha((\tilde{\delta}-\beta)^2 - \beta^2) - 2\beta(2 + (\beta - \tilde{\delta}))(1-\alpha)(2-\tilde{\delta})^d}{a_{max}(2-\tilde{\delta})^{2d}(2-\alpha)^2}.
$$

From London's Lemma [8] it follows that $\operatorname{Re}G(\mathbf{z})$ is positive definite $\forall \mathbf{z} \in \Theta_\beta$.

(b) By applying the Lidskii inequality (If $A, B \in \mathbb{C}^{d\times d}$ are Hermitian then $\lambda_{max}(A+B) \leqslant \lambda_{max}(A) + \lambda_{max}(B)$) we have that

$$
\begin{aligned}
\lambda_{max}(\operatorname{Re}\partial F(\mathbf{z})^T\partial F(\mathbf{z})) &\leqslant \lambda_{max}((I + R(\mathbf{y}) + R_r(\mathbf{w}))^T(I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\
&\quad - \lambda_{min}(R_i(\mathbf{w})^T R_i(\mathbf{w})) \\
&= \sigma^2_{max}(I + R(\mathbf{y}) + R_r(\mathbf{w})) - \sigma^2_{min}(R_i(\mathbf{w})) \\
&\leqslant (\sigma_{max}(I + R(\mathbf{y})) + \sigma_{max}(R_r(\mathbf{w})))^2 - \sigma^2_{max}(R_i(\mathbf{w})) \\
&\leqslant (\tilde{\delta} + \beta)^2.
\end{aligned} \tag{12}
$$

and

$$
\begin{aligned}
|\lambda_{max}(\operatorname{Im}\partial F(\mathbf{z})^T\partial F(\mathbf{z}))| &\leqslant |\lambda_{max}(R_i(\mathbf{w})^T(I + R(\mathbf{y}) + R_r(\mathbf{w})) \\
&\quad + (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w}))| \\
&\leqslant 2\beta(2 + (\beta - \tilde{\delta})).
\end{aligned} \tag{13}
$$

From inequalities (12) and (13), and Lemmas 3 and 4, we obtain

$$
\begin{aligned}
\lambda_{max}(\operatorname{Re}G(\mathbf{z})^{-1}) &\leqslant \frac{|\xi_R(\mathbf{z})|\lambda_{max}(\operatorname{Re}\partial F(\mathbf{z})^T\partial F(\mathbf{z})) + |\xi_I(\mathbf{z})||\lambda_{max}(\operatorname{Im}\partial F(\mathbf{z})^T\partial F(\mathbf{z}))|}{a_{min}|\xi(\mathbf{z})|^2} \\
&\leqslant D(\tilde{\delta}, \beta, d, a_{min}) < \infty
\end{aligned}
$$

where

$$
\begin{aligned}
D(\tilde{\delta}, \beta, d, a_{min}) := \frac{1}{a_{min}\tilde{\delta}^{2d}\alpha^2}\Big[&(2-\tilde{\delta})^d(2-\alpha)(\tilde{\delta}+\beta)^2 \\
&+2(2-\tilde{\delta})^d(1-\alpha)(\beta(2 + (\beta - \tilde{\delta}))\Big].
\end{aligned}
$$

(c) Similarly we can bound

$$
\begin{aligned}
\sigma_{max}(\operatorname{Im}\partial F(\mathbf{z})^T \partial F(\mathbf{z})) &\leqslant \sigma_{max}(R_i(\mathbf{w})^T(I + R(\mathbf{y}) + R_r(\mathbf{w})) \\
&+ (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w})) \\
&\leqslant 2\sigma_{max}(R_i(\mathbf{w}))\sigma_{max}(I + R(\mathbf{y}) + R_r(\mathbf{w})) \\
&\leqslant 2\beta(2 + (\beta - \tilde{\delta})).
\end{aligned} \tag{14}
$$

and

$$
\begin{aligned}
\sigma_{max}(\operatorname{Re}\partial F(\mathbf{z})^T \partial F(\mathbf{z})) &\leqslant \sigma_{max}((I + R(\mathbf{y}) + R_r(\mathbf{w}))^T(I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\
&+ \sigma_{max}(R_i(\mathbf{w})^T R_i(\mathbf{w})) \\
&= \sigma_{max}^2(I + R(\mathbf{y}) + R_r(\mathbf{w})) + \sigma_{max}^2(R_i(\mathbf{w})) \\
&\leqslant (\sigma_{max}(I + R(\mathbf{y})) + \sigma_{max}(R_r(\mathbf{w})))^2 + \sigma_{max}^2(R_i(\mathbf{w})) \\
&\leqslant ((2 - \tilde{\delta}) + \beta)^2 + \beta^2.
\end{aligned} \tag{15}
$$

From inequalities (14) and (15), and Lemmas 3 and 4 we obtain

$$
\begin{aligned}
\sigma_{max}(\operatorname{Im} G(\mathbf{z})^{-1}) &\leqslant \frac{\sigma_{max}(\xi_R(\mathbf{z}) \operatorname{Im}\partial F(\mathbf{z})^T \partial F(\mathbf{z}) - \xi_I(\mathbf{z}) \operatorname{Re}\partial F(\mathbf{z})^T \partial F(\mathbf{z}))}{a_{min}|\xi(\mathbf{z})|^2} \\
&\leqslant \frac{|\xi_R(\mathbf{z})|\sigma_{max}(\operatorname{Im}\partial F(\mathbf{z})^T \partial F(\mathbf{z})) + |\xi_I(\mathbf{z})|\sigma_{max}(\operatorname{Re}\partial F(\mathbf{z})^T \partial F(\mathbf{z}))}{a_{min}|\xi(\mathbf{z})|^2} \\
&= C(\tilde{\delta}, \beta, d, a_{min}) < \infty.
\end{aligned}
$$

where

$$
\begin{aligned}
C(\tilde{\delta}, \beta, d, a_{min}) := \frac{1}{a_{min}\tilde{\delta}^{2d}\alpha^2} \Big[&(2 - \tilde{\delta})^d (2 - \alpha) 2\beta(2 + (\beta - \tilde{\delta})) \\
&+ (2 - \tilde{\delta})^d (1 - \alpha)((2 - \tilde{\delta}) + \beta)^2 + \beta^2)\Big].
\end{aligned}
$$

□

Lemma 6. *Let* $G(\mathbf{z}) := (a \circ F)det(\partial F(\mathbf{z}))\partial F^{-1}(\mathbf{z})\partial F^{-T}(\mathbf{z})$ *then* $G(\mathbf{z})$ *is positive definite* $\forall \mathbf{z} \in \Theta_\beta$ *whenever*

$$
0 < \beta < min\{\tilde{\delta}\frac{\log(2 - \gamma)}{d + \log(2 - \gamma)}, \sqrt{1 + \tilde{\delta}^2/2} - 1\} \tag{16}
$$

where $\gamma := \frac{(2-\tilde{\delta})^d + 2\tilde{\delta}^d}{\tilde{\delta}^d + (2-\tilde{\delta})^d}$*. Furthermore*

$$
\lambda_{min}(\operatorname{Re} G(\mathbf{z})) \geqslant \varepsilon(\tilde{\delta}, \beta, d, a_{max}, a_{min}) > 0
$$

where

$$
\varepsilon(\tilde{\delta}, \beta, d, a_{max}, a_{min}) := \frac{1}{\left(1 + \left(\frac{C(\tilde{\delta},\beta,d,a_{min})}{B(\tilde{\delta},\beta,d,a_{max})}\right)^2\right) D(\tilde{\delta}, \beta, d, a_{min})}.
$$

Proof. From Lemma 5 $\operatorname{Re} G(\mathbf{z})$ is positive definite $\forall \mathbf{z} \in \Theta_\beta$, where β satisfies (16). It follows from the Lemma in [8] that $G(\mathbf{z}) = Q(I + i\Lambda)Q^*$, where Q is a non-singular matrix, $\Lambda := diag(\alpha_1, \dots, \alpha_d)$ and $\alpha_1, \dots, \alpha_d$ are real. Since $G(\mathbf{z})$ is symmetric then $\operatorname{Re} G(\mathbf{z}) = (1/2)(G(\mathbf{z}) + G(\mathbf{z})^*)$ and it is simple to see that $\operatorname{Re} G(\mathbf{z}) = QQ^*$. Thus we need to show that $\lambda_{min}(\operatorname{Re} G(\mathbf{z})) = \sigma^2_{min}(Q) \geqslant \varepsilon > 0$. Applying (b) in [8] we have that

$$G(\mathbf{z})^{-1} = (DQ^{-1})^*(I - i\Lambda)DQ^{-1}$$

where $D := diag((1+\alpha_1^2)^{-1/2}, \dots, (1+\alpha_d^2)^{-1/2}) = (I + \Lambda^2)^{-1/2}$. It follows that $\operatorname{Re} G(\mathbf{z})^{-1} = (DQ^{-1})^* DQ^{-1}$ for all $\mathbf{z} \in \Theta_\beta$,

$$\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1}) = \sigma^2_{max}(DQ^{-1}) \geqslant \sigma^2_{min}(D)\sigma^2_{max}(Q^{-1}) = \sigma^2_{min}(D)\sigma^{-2}_{min}(Q),$$

and therefore

$$\sigma^2_{min}(Q) \geqslant \frac{\sigma^2_{min}(D)}{|\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1})|} = \frac{\sigma_{min}((I+\Lambda^2)^{-1})}{|\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1})|} \geqslant \frac{(1+\sigma^2_{max}(\Lambda))^{-1}}{|\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1})|}$$

for all $\mathbf{z} \in \Theta_\beta$. Now, $\operatorname{Im} G(\mathbf{z})^{-1} = (DQ^{-1})^*(-\Lambda)DQ^{-1}$ and

$$\sigma_{max}(\operatorname{Im} G(\mathbf{z})^{-1}) \geqslant \sigma^2_{min}(DQ^{-1})\sigma_{max}(\Lambda).$$

Since $\operatorname{Re} G(\mathbf{z})^{-1} = (DQ^{-1})^* DQ^{-1}$ then $\lambda_{min}(\operatorname{Re} G(\mathbf{z})^{-1}) = \sigma^2_{min}(DQ^{-1})$ and

$$A(\tilde{\delta}, \beta, d, a_{max}, a_{min}) := \frac{\sigma_{max}(\operatorname{Im} G(\mathbf{z})^{-1})}{\lambda_{min}(\operatorname{Re} G(\mathbf{z})^{-1})} \geqslant \sigma_{max}(\Lambda).$$

It follows that

$$\lambda_{min}(\operatorname{Re} G(\mathbf{z})) \geqslant \frac{1}{(1+A^2)|\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1})|}. \tag{17}$$

From Lemma 5 (a) we have that $\lambda_{min}(\operatorname{Re} G(\mathbf{z})^{-1}) \geqslant B(\tilde{\delta}, \beta, d, a_{max})$. From Lemma 5 (c) we have that $\sigma_{max}(\operatorname{Im} G(\mathbf{z})^{-1}) \leqslant C(\tilde{\delta}, \beta, d, a_{min}) < \infty$. This implies $\sigma_{max}(\Lambda) \leqslant A(\tilde{\delta}, \beta, d, a_{max}, a_{min}) < \infty$. Finally from Lemma 5 (b) $\lambda_{max}(\operatorname{Re} G(\mathbf{z})^{-1}) \leqslant D(\tilde{\delta}, \beta, d, a_{min}) < \infty$. We conclude that

$$\lambda_{min}(\operatorname{Re} G(\mathbf{z})) \geqslant \varepsilon(\tilde{\delta}, \beta, d, a_{max}, a_{min}) > 0.$$

□

We are now ready to prove the main result of this section. For $n = 1, \dots, N$ consider the map $\Psi(s) : \Gamma_n \to H^1_0(\tilde{U})$ where

$$\Psi(s) := u(y_n(s), \hat{\mathbf{y}}_n, x),$$

for any arbitrary point $\hat{\mathbf{y}}_n \in \hat{\Gamma}_n$ where

$$\hat{\Gamma}_n := \left(\prod_{l=1,\dots,N, l \neq n} \Gamma_l \right).$$

Consider the extension of s into the complex plane as $z = s + iw$ in the region Θ_β along the n^{th} dimension. Now, for notational simplicity reorder $(y_1, \dots, y_N)$ such that $n = N$ and extend $\hat{\mathbf{y}}_n \to \hat{\mathbf{z}} \in \hat{\Theta}^n_\beta$, where $\hat{\Theta}^n_\beta := \Theta_\beta \cap \mathbb{C}^{N-1}$. Then $\Psi(s)$ has a natural extension to the complex plane as $\Psi(z) := u(z, \hat{\mathbf{z}}, x)$ for all $\hat{\mathbf{z}} \in \hat{\Theta}^n_\beta \subset \Theta_\beta$.

Theorem 1. *Let* $0 < \tilde{\delta} < 1$ *then* $u(\mathbf{z})$ *is holomorphic in* Θ_β (8) *if*

$$\beta < min\{\tilde{\delta}\frac{\log(2-\gamma)}{d + \log(2-\gamma)}, \sqrt{1 + \tilde{\delta}^2/2} - 1\}$$

where $\gamma := \frac{(2-\tilde{\delta})^d + 2\tilde{\delta}^d}{\tilde{\delta}^d + (2-\tilde{\delta})^d}$. *Moreover, the following estimate holds:*

$$\|u(\mathbf{z})\|_{H_0^1(\tilde{U})} \leqslant \frac{C_P(\tilde{U})\|f\|_{L^2(\tilde{U})}}{\varepsilon(\tilde{\delta},\beta,d,a_{max},a_{min})} < \infty, \tag{18}$$

with $\varepsilon(\tilde{\delta},\beta,d,a_{max},a_{min})$ *defined in Lemmas 5 and 6.*

Proof. The strategy for this proof is show that $\Psi(s)$ admits an analytic extension into the complex plain for each dimension separately (for $n = 1,\dots,N$) and then apply Hartog's Theorem (Chap1, p32, [9]) and Osgood's Lemma (Chap 1, p 2, [10]) to show that it extends to the entire domain Θ_β. First, since $\beta < \tilde{\delta}$ the series

$$\partial F^{-1}(\mathbf{z}) = (I + R(\mathbf{z}))^{-1} = I + \sum_{k=1}^{\infty} R(\mathbf{z})^k$$

is convergent $\forall \mathbf{z} \in \Theta_\beta$. It follows that each entry of $\partial F(\mathbf{z})^{-1}$ is analytic for all $\mathbf{z} \in \Theta_\beta$. From Lemma 4 it follows that the entries of $G(\mathbf{z})$ are analytic for all $\mathbf{z} \in \Theta_\beta$.

Let $\Psi = [\Psi_R,\ \Psi_I]^T$, where $\Psi_R = Re\ \Psi(z)$ and $\Psi_I = Im\ \Psi(z)$. Then Ψ solves (in the weak sense) the problem

$$-\nabla \cdot \hat{G}\nabla\Psi = \hat{f}, \tag{19}$$

where

$$\hat{G} := \begin{pmatrix} G_R & -G_I \\ G_I & G_R \end{pmatrix},\ \ \hat{f} := \begin{pmatrix} f_R \\ f_I \end{pmatrix},$$

$G_R := \mathrm{Re}(G)$, $G_I := \mathrm{Im}(G)$, $f_R := \mathrm{Re}\,\tilde{f}$ and $f_I = \mathrm{Im}\,\tilde{f}$. Note that $\tilde{f}$ refers to rhs of the weak formulation i.e. $\tilde{l}(z;v)$ for all $v \in H_0^1(U)$.

The system of equations (19) has a unique solution if G_R is positive definite ($\lambda_{min}(G_R(\mathbf{z})) > 0$). From Lemma (6) this condition is satisfied if $\mathbf{z} \in \Theta_\beta$.

To show that $\Psi(z) : \mathbb{C} \to H_0^1(\tilde{U})$ is holomorphic in $\mathbb{C}$ for $n = 1,\dots,N$ the strategy is to show that the Cauchy-Riemann conditions are satisfied, but first we have to show that the derivatives $\partial_s\Psi$ and $\partial_w\Psi$ exist. Now, differentiating (19) with respect to $s = \mathrm{Re}\,z$ and $w = \mathrm{Im}\,z$ we obtain

$$\begin{aligned}
-(\nabla\cdot G_R\nabla\partial_s\Psi_R(z) - \nabla\cdot G_I\nabla\partial_s\Psi_I(z)) &= \nabla\cdot\partial_s G_R\nabla\Psi_R(z) - \nabla\cdot\partial_s G_I\nabla\Psi_I(z) \\
&+ \partial_s f_R(z) \\
-(\nabla\cdot G_I\nabla\partial_s\Psi_R(z) + \nabla\cdot G_R\nabla\partial_s\Psi_I(z)) &= \nabla\cdot\partial_s G_I\nabla\Psi_R(z) + \nabla\cdot\partial_s G_R\nabla\Psi_I(z) \\
&+ \partial_s f_I(z) \\
-(\nabla\cdot G_R\nabla\partial_w\Psi_R(z) - \nabla\cdot G_I\nabla\partial_w\Psi_I(z)) &= \nabla\cdot\partial_w G_R\nabla\Psi_R(z) - \nabla\cdot\partial_w G_I\nabla\Psi_I(z) \\
&+ \partial_w f_R(z) \\
-(\nabla\cdot G_I\nabla\partial_w\Psi_R(z) + \nabla\cdot G_R\nabla\partial_w\Psi_I(z)) &= \nabla\cdot\partial_w G_I\nabla\Psi_R(z) + \nabla\cdot\partial_w G_R\nabla\Psi_I(z) \\
&+ \partial_w f_I(z)
\end{aligned} \tag{20}$$

By the Lax-Milgram theorem the derivatives $\partial_s\Psi$ and $\partial_w\Psi$ exist and have a unique solution whenever $z_n \in \Theta_\beta$ and $\hat{\mathbf{z}} \in \hat{\Theta}_\beta^n$. The second step is now to show that the Cauchy-Riemann conditions are satisfied.

Let $P(z) := \partial_s\Psi_R(z) - \partial_w\Psi_I(z)$ and $Q(z) := \partial_w\Psi_R(z) + \partial_s\Psi_I(z)$. To show analyticity we now have to analyze for what region in the complex plane $P(z) = 0$ and $Q(z) = 0$. By taking linear combinations of eqns (20) we obtain

$$
\begin{aligned}
-\nabla \cdot (G_R \nabla P - G_I \nabla Q) &= \nabla \cdot ((\partial_s G_R - \partial_w G_I)\nabla \Psi_R - (\partial_w G_R + \partial_s G_I)\nabla \Psi_I \\
&+ \partial_s f_R - \partial_w f_I \\
-\nabla \cdot (G_I \nabla P + G_R \nabla Q) &= \nabla \cdot ((\partial_w G_R + \partial_s G_I)\nabla \Psi_R - (\partial_s G_R - \partial_w G_I)\nabla \Psi_I \\
&+ \partial_s f_I + \partial_w f_R
\end{aligned} \tag{21}
$$

We now need to show that $G(z)$ and $\hat{f}(z)$ satisfies the Riemann-Cauchy conditions so that the right hand side becomes zero.

From Assumption 5 (c) we have that $f \circ F(z)$ and $\mathbf{w} \circ F$ can be analytically extended in $\mathbb{C}$ thus $l(z, \hat{\mathbf{z}}; v)$ is holomorphic for all $z \in \Theta_\beta$, $\hat{\mathbf{z}} \in \hat{\Theta}^n_\beta$, and $v \in H^1_0(U)$. Now, recall that $G(\mathbf{z})$ is analytic if $\mathbf{z} \in \Theta_\beta$. Thus equations (21) have a unique solution $P(z) = Q(z) = 0$ for all $z \in \Theta_\beta$ and $\hat{\mathbf{z}} \in \hat{\Theta}^n_\beta$. From the Looman-Menchoff theorem $\Psi(z)$ is holomorphic for all $z \in \Theta_\beta$ and $\hat{\mathbf{z}} \in \hat{\Theta}^n_\beta$.

We can now extend the analyticity of the solution $u(\mathbf{z})$ to the entire domain Θ_β. Repeat the analytic extension of $u(y_n, \hat{\mathbf{y}}_n, x)$ for $n = 1, \dots, N$. Since each variable $u(y_n, \hat{\mathbf{y}}_n, x)$ has been extended into the complex plane for $z \in \Theta_\beta$ and $\hat{\mathbf{z}} \in \hat{\Theta}^n_\beta$ from Hartog's Theorem it follows that $\Psi(\mathbf{z})$ is continuous in Θ_β. From Osgood's Lemma it follows that $\Psi(\mathbf{z})$ is holomorphic for all $\mathbf{z} \in \Theta_\beta$.

The last step is to show the inequality (18). First, multiply (19) by $\Psi(\mathbf{z})^T$ and integrate over $\tilde{U}$ to obtain

$$
\int_D \nabla \Psi_R^T(\mathbf{z}) G_R(\mathbf{z}) \nabla \Psi_R(\mathbf{z}) + \nabla \Psi_I^T(\mathbf{z}) G_R(\mathbf{z}) \nabla \Psi_I(\mathbf{z}) \leqslant \|f(\mathbf{z})\|_{L^2(\tilde{U})} \|\Psi(\mathbf{z})\|_{L^2(\tilde{U})},
$$

thus

$$
\lambda_{min}(G_R(\mathbf{z})) \|\nabla \Psi(\mathbf{z})\|^2_{L^2(\tilde{U})} \leqslant \|f(\mathbf{z})\|_{L^2(\tilde{U})} \|\Psi(\mathbf{z})\|_{L^2(\tilde{U})}.
$$

Applying Poincaré inequality and Lemma 6 we obtain the result. □

4. Stochastic Collocation

We seek to efficiently approximate the mean and variance of the QoI of the form (4). More specifically we seek a numerical approximation to the exact moments of the QoI in a finite dimensional subspace $V_{p,h}$ based on a tensor product structure, where the following hold:

- $H_h(U) \subset H^1_0(U)$ is a standard finite element space of dimension N_h, which contains continuous piecewise polynomials defined on regular triangulations $\mathcal{T}_h$ that have a maximum mesh spacing parameter $h > 0$.
- $\mathcal{P}_p(\Gamma) \subset L^2_\rho(\Gamma)$ is the span of tensor product polynomials of degree at most $p = (p_1, \dots, p_N)$; i.e., $\mathcal{P}_p(\Gamma) = \bigotimes_{n=1}^N \mathcal{P}_{p_n}(\Gamma_n)$ with

  $$\mathcal{P}_{p_n}(\Gamma_n) = \text{span}(y_n^m,\, m = 0, \dots, p_n), \quad n = 1, \dots, N.$$

  Hence the dimension of $\mathcal{P}_p$ is $N_p = \prod_{n=1}^N (p_n + 1)$.
- $u_h : \Gamma \to H_h(U)$ is the semidiscrete approximation that is obtained by projecting the solution of (3) onto the subspace $H_h(U)$, for each $\mathbf{y} \in \Gamma$, i.e.,

$$
\int_U a \circ F [\nabla u_h(\mathbf{y})]^T G(\mathbf{y}) \nabla v_h \, dx = \int_U (f \circ F - L\mathbf{w}) v_h \, dx \quad \forall v_h \in H_h(U), \tag{22}
$$

for a.e. $\mathbf{y} \in \Gamma$. Denote $\pi_h : H^1_0(U) \to H_h(U)$ as the finite element operator s.t. if $u \in H^1_0(U)$ then $u_h := \pi_h u$ and

$$\|u - \pi_h u\|_{H^1_0(U)} \leqslant C_\pi \min_{v \in H_h(U)} \|u - v\|_{H^1_0(U)} \leqslant h^r C(r, u). \tag{23}$$

The constant $r \in \mathbb{N}$ will depend on the regularity of u and the polynomial order of the finite element space H_h. Denote $C_\Gamma(r) := \sup_{\mathbf{y} \in \Gamma} C(r, u(\mathbf{y}))$.

- Similarly, $\varphi_h := \pi_h \varphi$ is the semi-discrete approximation of the influence function. For each $\mathbf{y} \in \Gamma$, i.e.,

$$\int_U a \circ F[\nabla v_h(\mathbf{y})]^T G(\mathbf{y}) \nabla \varphi_h \, dx = Q(v_h) \quad \forall v_h \in H_h(U). \tag{24}$$

Remark 4. *Note that for the sake of simplicity we ignore quadrature errors and assume that the integrals* (22) *and* (24) *are computed exactly.*

The next step consists in collocating $Q_h(u_h(\mathbf{y}))$ with respect to Γ. To this end, we first introduce an auxiliary probability density function $\hat{\rho} : \Gamma \to \mathbb{R}^+$ that can be seen as the joint probability of N independent random variables; i.e., it factorizes as

$$\hat{\rho}(\mathbf{y}) = \prod_{n=1}^{N} \hat{\rho}_n(y_n) \;\; \forall \mathbf{y} \in \Gamma, \qquad \text{and is such that} \quad \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma)} < \infty. \tag{25}$$

For each dimension $n = 1, \ldots, N$, let y_{n,k_n}, $1 \leqslant k_n \leqslant p_n + 1$, be the $p_n + 1$ roots of the orthogonal polynomial q_{p_n+1} with respect to the weight $\hat{\rho}_n$, which then satisfies $\int_{\Gamma_n} q_{p_n+1}(\mathbf{y}) v(\mathbf{y}) \hat{\rho}_n(\mathbf{y}) dy = 0$ for all $v \in \mathcal{P}_{p_n}(\Gamma_n)$.

Standard choices for $\hat{\rho}$, such as constant, Gaussian, etc., lead to well-known roots of the polynomial q_{p_n+1}, which are tabulated to full accuracy and do not need to be computed. Note, that for the case of Clenshaw-Curtis abscissas the collocation points are chosen as extrema of Chebyshev polynomials.

To any vector of indexes $[k_1, \ldots, k_N]$ we associate the global index

$$k = k_1 + p_1(k_2 - 1) + p_1 p_2 (k_3 - 1) + \cdots$$

and we denote by y_k the point $y_k = [y_{1,k_1}, y_{2,k_2}, \ldots, y_{N,k_N}] \in \Gamma$. We also introduce, for each $n = 1, 2, \ldots, N$, the Lagrange basis $\{l_{n,j}\}_{j=1}^{p_n+1}$ of the space $\mathcal{P}_{p_n}$,

$$l_{n,j} \in \mathcal{P}_{p_n}(\Gamma_n), \qquad l_{n,j}(y_{n,k}) = \tilde{\delta}_{jk}, \quad j, k = 1, \ldots, p_n + 1,$$

where $\tilde{\delta}_{jk}$ is the Kronecker symbol, and we set $l_k(\mathbf{y}) = \prod_{n=1}^{N} l_{n,k_n}(y_n)$. Now, let $\mathcal{I}_p : C^0(\Gamma) \to \mathcal{P}_p(\Gamma)$, such that

$$\mathcal{I}_p v(\mathbf{y}) = \sum_{k=1}^{N_p} v(y_k) l_k(\mathbf{y}) \qquad \forall v \in C^0(\Gamma).$$

Thus for any $\mathbf{y} \in \Gamma$ we can write the Lagrange approximation of the QoI ($Q_h(\mathbf{y})$):

$$Q_{h,p}(\mathbf{y}) := \mathcal{I}_p B(\mathbf{y}; u_h(\mathbf{y}), \varphi_h(\mathbf{y}))$$

Remark 5. *For any continuous function $g : \Gamma \to \mathbb{R}$ we introduce the Gauss quadrature formula $\mathbb{E}^p_{\hat{\rho}}[g]$ approximating the integral $\int_\Gamma g(\mathbf{y})\hat{\rho}(\mathbf{y})\, d\mathbf{y}$ as*

$$\mathbb{E}^p_{\hat{\rho}}[g] = \sum_{k=1}^{N_p} \omega_k g(y_k), \quad \omega_k = \prod_{n=1}^{N} \omega_{k_n}, \quad \omega_{k_n} = \int_{\Gamma_n} l^2_{k_n}(y)\hat{\rho}_n(y)\, dy. \tag{26}$$

In the case $\rho/\hat{\rho}$ is a smooth function we can use directly (26) *to approximate the mean value or the variance of Q_h as*

$$\mathbb{E}_h[Q_h] := \mathbb{E}^p_{\hat{\rho}}\left[\frac{\rho}{\hat{\rho}} Q_{h,p}\right], \text{ and } \mathrm{var}_h(Q_h) := \mathbb{E}^p_{\hat{\rho}}\left[\frac{\rho}{\hat{\rho}} Q^2_{h,p}\right] - \mathbb{E}^p_{\hat{\rho}}\left[\frac{\rho}{\hat{\rho}} Q_{h,p}\right]^2.$$

Otherwise, $\mathbb{E}[Q_h]$ and $\mathrm{var}_h(Q_h)$ should be computed with a suitable quadrature formula that takes into account eventual discontinuities or singularities of $\rho/\hat{\rho}$. However, to simplify the error analysis presentation in Section 5, we shall assume that the quadrature scheme for the expectation to be exact.

4.1. **Sparse Grid Approximation.** Recall that the dimension of $\mathcal{P}_p$ increases as $\prod_{n=1}^N (p_n + 1)$. This has the consequence that even for a relatively small dimension N the accurate computation of the mean and variance of the QoI with a tensor product grid becomes intractable. However, if the stochastic integral is highly regular with respect to the random variables, the application of Smolyak sparse grids is well suited. We present here a generalization of the classical Smolyak construction (see e.g. [11, 12]) to build a multivariate polynomial approximation on a sparse grid. See [13] for details.

Let $\mathcal{I}^{m(i)}_n : C^0(\Gamma_n) \to \mathcal{P}_{m(i)-1}(\Gamma_n)$ be the 1D interpolant as previously introduced. Here $i \geqslant 1$ denotes the level of approximation and $m(i)$ the number of collocation points used to build the interpolation at level i, with the requirement that $m(1) = 1$ and $m(i) < m(i+1)$ for $i \geqslant 1$. In addition, let $m(0) = 0$ and $\mathcal{I}^{m(0)}_n = 0$. Further, we introduce the difference operators

$$\Delta^{m(i)}_n := \mathcal{I}^{m(i)}_n - \mathcal{I}^{m(i-1)}_n.$$

Given an integer $w \geqslant 0$ called the approximation level and a multi-index $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}^N_+$, we introduce a function $g : \mathbb{N}^N_+ \to \mathbb{N}$ strictly increasing in each argument and define a sparse grid approximation of Q_h

$$\mathcal{S}^{m,g}_w[Q_h] = \sum_{\mathbf{i} \in \mathbb{N}^N_+ : g(\mathbf{i}) \leqslant w} \bigotimes_{n=1}^{N} \Delta^{m(i_n)}_n(Q_h) \tag{27}$$

or equivalently written as

$$\mathcal{S}^{m,g}_w[Q_h] = \sum_{\mathbf{i} \in \mathbb{N}^N_+ : g(\mathbf{i}) \leqslant w} c(\mathbf{i}) \bigotimes_{n=1}^{N} \mathcal{I}^{m(i_n)}_n(Q_h), \qquad \text{with } c(\mathbf{i}) = \sum_{\substack{\mathbf{j} \in \{0,1\}^N : \\ g(\mathbf{i}+\mathbf{j}) \leqslant w}} (-1)^{|\mathbf{j}|}. \tag{28}$$

From the previous expression, we see that the sparse grid approximation is obtained as a linear combination of full tensor product interpolations. However, the constraint $g(\mathbf{i}) \leqslant w$ in (28) is typically chosen so as to forbid the use of tensor grids of high degree in all directions at the same time.

Let $\mathbf{m}(\mathbf{i}) = (m(i_1), \ldots, m(i_N))$ and consider the set of polynomial multi-degrees

$$\Lambda^{m,g}(w) = \{\mathbf{p} \in \mathbb{N}^N, \;\; g(\mathbf{m}^{-1}(\mathbf{p} + \mathbf{1})) \leqslant w\}.$$

Denote by $\mathbb{P}_{\Lambda^{m,g}(w)}(\Gamma)$ the corresponding multivariate polynomial space spanned by the monomials with multi-degree in $\Lambda^{m,g}(w)$, i.e.

$$\mathbb{P}_{\Lambda^{m,g}(w)}(\Gamma) = span\left\{\prod_{n=1}^{N} y_n^{p_n}, \text{ with } \mathbf{p} \in \Lambda^{m,g}(w)\right\}.$$

The following result proved in [13], states that the sparse approximation formula $\mathcal{S}_w^{m,g}$ is exact in $\mathbb{P}_{\Lambda^{m,g}(w)}(\Gamma)$:

Proposition 1.

a) *For any* $f \in C^0(\Gamma; V)$, *we have* $\mathcal{S}_w^{m,g}[f] \in \mathbb{P}_{\Lambda^{m,g}(w)} \otimes V$.
b) *Moreover,* $\mathcal{S}_w^{m,g}[v] = v, \ \forall v \in \mathbb{P}_{\Lambda^{m,g}(w)} \otimes V$.

Here V denotes a Banach space defined on U and

$$C^0(\Gamma; V) := \{v : \Gamma \times U \to V \text{ is continuous on } \Gamma \text{ and } \max_{y\in\Gamma} \|v(y)\|_V < \infty\}.$$

We recall that the most typical choice of m and g is given by (see [11, 12])

$$m(i) = \begin{cases} 1, & \text{for } i = 1 \\ 2^{i-1}+1, & \text{for } i > 1 \end{cases} \quad \text{and} \quad g(\mathbf{i}) = \sum_{n=1}^{N}(i_n - 1).$$

This choice of m, combined with the choice of Clenshaw-Curtis interpolation points (extrema of Chebyshev polynomials) leads to nested sequences of one dimensional interpolation formulas and a sparse grid with a highly reduced number of points compared to the corresponding tensor grid. In Table 1 different choices of $g(\mathbf{i})$ are given (see [13]).

Approx. space	sparse grid: m, g	polynomial space: $\Lambda(w)$
Tensor Product Product (TP)	$m(i) = i$ $g(\mathbf{i}) = \max_n(i_n - 1) \leqslant w$	$\{\mathbf{p} \in \mathbb{N}^N : \ \max_n p_n \leqslant w\}$
Total Degree (TD)	$m(i) = i$ $g(\mathbf{i}) = \sum_n(i_n - 1) \leqslant w$	$\{\mathbf{p} \in \mathbb{N}^N : \ \sum_n p_n \leqslant w\}$
Hyperbolic Cross (HC)	$m(i) = i$ $g(\mathbf{i}) = \prod_n(i_n) \leqslant w+1$	$\{\mathbf{p} \in \mathbb{N}^N : \ \prod_n(p_n+1) \leqslant w+1\}$
Smolyak (SM)	$m(i) = \begin{cases} 2^{i-1}+1, \ i > 1 \\ 1, \ i = 1 \end{cases}$ $g(\mathbf{i}) = \sum_n(i_n - 1) \leqslant w$	$\{\mathbf{p} \in \mathbb{N}^N : \ \sum_n f(p_n) \leqslant w\}$ $f(p) = \begin{cases} 0, \ p = 0 \\ 1, \ p = 1 \\ \lceil \log_2(p) \rceil, \ p \geqslant 2 \end{cases}$

TABLE 1. Sparse approximation formulas and corresponding set of polynomial multi-degrees used for approximation.

It is also straightforward to build related anisotropic sparse approximation formulas by making the function g to act differently on the input random variables y_n. Anisotropic sparse stochastic collocation [14] combines the advantages of isotropic sparse collocation with those of anisotropic full tensor product collocation.

The mean term $\mathbb{E}[Q_h]$ is approximated as

$$\mathbb{E}[\mathcal{S}_w^{m,g} Q_h] = \mathbb{E}_{\hat{\rho}}[\mathcal{S}_w^{m,g} Q_h \frac{\rho}{\hat{\rho}}], \tag{29}$$

where $v \in L^1_\rho(\Gamma)$

$$\mathbb{E}_{\hat{\rho}}[v] := \int_\Gamma v \hat{\rho}(\mathbf{y}) \; d\mathbf{y}$$

and similarly the variance var$[Q]$ is approximated as

$$\mathrm{var}_h[Q_h] = \mathbb{E}[(\mathcal{S}_w^{m,g}[Q_h])^2] - \mathbb{E}[\mathcal{S}_w^{m,g}[Q_h]]^2 = \mathbb{E}_{\hat{\rho}}[(\mathcal{S}_w^{m,g}[Q_h])^2 \frac{\rho}{\hat{\rho}}] - \mathbb{E}_{\hat{\rho}}[\mathcal{S}_w^{m,g}[Q_h] \frac{\rho}{\hat{\rho}}]^2. \tag{30}$$

5. Error Analysis

In this section we derive error estimates of the mean and variance with respect to (i) the finite element approximation, (ii) the sparse grid approximation and (iii) truncating the stochastic model to the first N_s dimensions.

For notational simplicity we split the Jacobian as follows

$$\partial F(x,\omega) = I + \sum_{l=1}^{N_s} B_l(x)\sqrt{\lambda_l} Y_l(\omega) + \sum_{l=N_s+1}^{N} B_l(x)\sqrt{\lambda_l} Y_l(\omega). \tag{31}$$

Furthermore, let $\Gamma_s := [-1,1]^{N_s}$, $\Gamma_f := [-1,1]^{N-N_s}$, then the domain $\Gamma = \Gamma_s \times \Gamma_f$. We now refer to $Q(\mathbf{y}_s)$ as $Q(\mathbf{y})$ restricted to the stochastic domain Γ_s and similarly for $G(\mathbf{y}_s)$. It is clear also that $Q(\mathbf{y}_s,\mathbf{y}_f) = Q(\mathbf{y})$ and $G(\mathbf{y}_s,\mathbf{y}_f) = G(\mathbf{y})$ for all $\mathbf{y} \in \Gamma_s \times \Gamma_f$, $\mathbf{y}_s \in \Gamma_s$, and $\mathbf{y}_f \in \Gamma_f$.

Now that we have established notation, we are interested in deriving estimates for the variance ($|var[Q(\mathbf{y}_s,\mathbf{y}_f)] - var[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]|$) and mean ($|\mathbb{E}\,[Q(\mathbf{y}_s,\mathbf{y}_f)] - \mathbb{E}\,[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]\,|$) errors. First observe that

$$\begin{aligned}
|var[Q(\mathbf{y}_s,\mathbf{y}_f)] - var[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]| \leqslant & |var[Q(\mathbf{y}_s,\mathbf{y}_f)] - var[Q(\mathbf{y}_s)]| \\
& + |var[Q(\mathbf{y}_s)] - var[Q_h(\mathbf{y}_s)]| \\
& + |var[Q_h(\mathbf{y}_s)] - var[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]|.
\end{aligned}$$

Let us analyze the first term. By applying the Cauchy-Schwartz inequality we have that

$$\mathbb{E}\left[Q(\mathbf{y}_s,\mathbf{y}_f)^2 - Q(\mathbf{y}_s)^2\right] \leqslant \|Q(\mathbf{y}_s,\mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)} \|Q(\mathbf{y}_s,\mathbf{y}_f) + Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)},$$

and

$$\begin{aligned}
|\mathbb{E}\,[Q(\mathbf{y}_s,\mathbf{y}_f)]^2 - \mathbb{E}\,[Q(\mathbf{y}_s)]^2\,| &= |\mathbb{E}\,[Q(\mathbf{y}_s,\mathbf{y}_f) - Q(\mathbf{y}_s)]\,\mathbb{E}\,[Q(\mathbf{y}_s,\mathbf{y}_f) + Q(\mathbf{y}_s)]\,| \\
&\leqslant \|Q(\mathbf{y}_s,\mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)} \|Q(\mathbf{y}_s,\mathbf{y}_f) + Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)}.
\end{aligned}$$

Therefore

$$|var[Q(\mathbf{y}_s,\mathbf{y}_f)] - var[Q(\mathbf{y}_s)]| \leqslant C_T \|Q(\mathbf{y}_s,\mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)}$$

for some positive constant $C_T \in \mathbb{R}^+$. It is not hard to show that $|var[Q(\mathbf{y}_s,\mathbf{y}_f)] - var[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]|$ and $|\mathbb{E}\,[Q(\mathbf{y}_s,\mathbf{y}_f)] - \mathbb{E}\,[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]|$ are less or equal to

$$C_T \underbrace{\|Q(\mathbf{y}_s,\mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)}}_{\text{Truncation (I)}} + C_{FE} \underbrace{\|Q(\mathbf{y}_s) - Q_h(\mathbf{y}_s)\|_{L^2_P(\Gamma_s)}}_{\text{Finite Element (II)}}$$

$$+C_{SG} \underbrace{\|Q_h(\mathbf{y}_s) - \mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]\|_{L^2_P(\Gamma_s)}}_{\text{Sparse Grid (III)}}.$$

for some positive constants C_T, C_{FE} and C_{SG}. We now study the error contributions from (I), (II) and (III).

5.1. **Truncation Error (I).** Given that $Q : H_0^1(U) \to \mathbb{R}$ is a bounded linear functional then for any realization of $\varphi(\mathbf{y}_s, \mathbf{y}_f)$ we have that

$$\begin{aligned}|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)| &= |B(\mathbf{y}_s, \mathbf{y}_f; \varphi(\mathbf{y}_s, \mathbf{y}_f), u(\mathbf{y}_s, \mathbf{y}_f) - u(\mathbf{y}_s))| \\ &\leqslant a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min}\|\varphi(\mathbf{y}_s, \mathbf{y}_f)\|_{H^1_0(U)}\|u(\mathbf{y}_s, \mathbf{y}_f) - u(\mathbf{y}_s)\|_{H^1_0(U)}.\end{aligned}$$

Following a similar argument as the proof from Lemma 2 we have that

$$\|\varphi(\mathbf{y}_s, \mathbf{y}_f)\|_{H^1_0(U)} \leqslant \frac{\|q \circ F\|_{L^2(U)} C_P(U)\mathbb{F}^{d+2}_{max}}{a_{min}\mathbb{F}^d_{min}}$$

a.s., where q is defined in eqn (5). Thus

$$\|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_P(\Gamma)} \leqslant C_{TR}\|u(\mathbf{y}_s, \mathbf{y}_f) - u(\mathbf{y}_s)\|_{L^2(\Gamma; H^1_0(U))},$$

where $C_{TR} := a_{max}a^{-1}_{min}\mathbb{F}^{2d+2}_{max}\mathbb{F}^{-d-2}_{min}\|q \circ F\|_{L^2(U)}C_P(U)$. We now seek control on the error term $e := \|u(\mathbf{y}_s, \mathbf{y}_f) - u(\mathbf{y}_s)\|_{L^2_P(\Gamma; H^1_0(U))}$. First we establish some notation and definitions. From Section 3 we have shown that the solution u of Problem 3 varies continuously with respect to $y \in \Gamma$. More precisely, recall that if V is a Banach space defined on U and

$$C^0(\Gamma; V) := \{v : \Gamma \times U \to V \text{ is continuous on } \Gamma \text{ and } \max_{y \in \Gamma} \|v(y)\|_V < \infty\},$$

then $u \in C^0(\Gamma, H^1_0(U))$. Furthermore, let

$$L^2_\rho(\Gamma; V) := \{v : \Gamma \times U \to V \text{ is strongly measurable and } \int_\Gamma \|v\|^2_V\, \rho(\mathbf{y})\, d\mathbf{y} < \infty\}.$$

From Assumption 4 we have that $u \in C^0(\Gamma; H^1_0(U)) \subset L^2_\rho(\Gamma; H^1_0(U))$, thus u satisfies the following variational problem

$$\mathcal{A}(u, v) := E[B(\mathbf{y}_s, \mathbf{y}_f; u, v)] = E[\tilde{l}(\mathbf{y}_s, \mathbf{y}_f; v)] \;\; \forall v \in L^2_\rho(\Gamma; H^1_0(U)).$$

The following lemma will be useful in deriving error estimates.

Lemma 7. *For all $w, v \in L^2_\rho(\Gamma; H^1_0(U))$ we have that*

$$|\mathcal{A}(w, v)| \leqslant a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min}\|w\|_{L^2_\rho(\Gamma; H^1_0(U))}\|v\|_{L^2_\rho(\Gamma; H^1_0(U))}.$$

Proof.

$$\begin{aligned}|\mathcal{A}(w, v)| &\leqslant \sup_{\mathbf{y} \in \Gamma}\lambda_{max}(G(\mathbf{y}))\mathbb{E}\left[\int_U |\nabla w^T \nabla v|\right] \\ &\leqslant a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min}\mathbb{E}[\|\nabla w\|_{L^2(U)}\|\nabla v\|_{L^2(U)}] \\ &\leqslant a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min}\|w\|_{L^2_\rho(\Gamma; H^1_0(U))}\|v\|_{L^2_\rho(\Gamma; H^1_0(U))}\end{aligned}$$

□

We can now derive the truncation error (I).

Theorem 2. *Let u be the solution to the bilinear Problem 3 that satisfies Assumptions 1, 2, 3, 4 and 5. Furthermore, let* $B_{\mathbb{T}} := \sup_{x \in U} \sum_{i=N_s+1}^{N} \sqrt{\lambda_i} \|B_l(x)\|$ *then*

$$\|u(\mathbf{y}_s, \mathbf{y}_f) - u(\mathbf{y}_s)\|_{L^2_\rho(\Gamma;H^1_0(U))} \leqslant \frac{C_P(U)^2}{a_{min}\mathbb{F}^d_{min}\mathbb{F}^{-2}_{max}} \sup_{x \in U, \mathbf{y} \in \Gamma} \|G(\mathbf{y}_s, \mathbf{y}_f) - G(\mathbf{y}_s)\| * \|u(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2_\rho(\Gamma;H^1_0(U))},$$

where

$$\sup_{x \in U, \mathbf{y} \in \Gamma} \|G(\mathbf{y}_s, \mathbf{y}_f) - G(\mathbf{y}_s)\| \leqslant \|a\|_{W^{1,\infty}(U)} |\hat{v}| \sum_{i=N_s+1}^{N} \sqrt{\lambda_i} \|b_i(x)\|_{L^\infty(U)} + a_{max} B_{\mathbb{T}} H(\mathbb{F}_{max}, \mathbb{F}_{min}, \tilde{\delta}, d)$$

and $H(\mathbb{F}_{max}, \mathbb{F}_{min}, \tilde{\delta}, d) := \mathbb{F}^{d-1}_{max}\mathbb{F}^{-3}_{min}(\mathbb{F}_{max}(2 + \mathbb{F}^{-1}_{min}(1 - \tilde{\delta})) + \mathbb{F}^{-1}_{min} d)$.

Proof. We follow a similar strategy as in [16, 17] to compute the bounds for the truncation of the stochastic variables to Γ_s. Consider the solution to Problem 3 $u_{N_s} \in C^0(\Gamma_s; H^1_0(U)) \subset L^2_\rho(\Gamma; H^1_0(U))$ where the matrix of coefficients $G(\mathbf{y}_s)$ depends only on the variables $Y_1, \dots, Y_{N_s}$,

$$\mathbb{E}[B(\mathbf{y}_s; u_{N_s}, v)] = \mathbb{E}[\tilde{l}(\mathbf{y}_s; v)] \;\; \forall v \in L^2_\rho(\Gamma_s; H^1_0(U)).$$

Furthermore the variational form is still valid $\forall v \in L^2_\rho(\Gamma; H^1_0(U))$ i.e. for all $\mathbf{y}_s \in \Gamma_s$

$$\mathcal{A}_{N_s}(u_{N_s}, v) := \mathbb{E}\left[B(\mathbf{y}_s; u_{N_s}, v)\right] = \mathbb{E}\left[\tilde{l}(\mathbf{y}_s; v)\right] \;\; \forall v \in L^2_\rho(\Gamma; H^1_0(U)).$$

Now, Observe that $\forall v \in L^2_\rho(\Gamma; H^1_0(U))$ we have that

$$\begin{aligned}
\mathcal{A}_{N_s}(v, v) &\geqslant \inf_{\mathbf{y}_s \in \Gamma_s} \lambda_{min}(G(\mathbf{y}_s)) \mathbb{E}[\|\nabla v\|^2_{L^2(U)}] \\
&\geqslant a_{min}\mathbb{F}^d_{min}\mathbb{F}^{-2}_{max} \mathbb{E}[\|\nabla v\|^2_{L^2(U)}] \\
&\geqslant a_{min}\mathbb{F}^d_{min}\mathbb{F}^{-2}_{max} C_P(U)^{-2} \|v\|^2_{L^2(\Gamma;H^1_0(U))}.
\end{aligned}$$

By adapting the proof from Strang's Lemma and applying Lemma 7 we have that for all $v \in L^2_\rho(\Gamma; H^1_0(U))$

$$\begin{aligned}
\|u_{N_s} - v\|^2_{L^2_\rho(\Gamma;H^1_0(U))} &\leqslant \mathcal{C}_1(\mathcal{A}_{N_s}(u_{N_s} - v, u_{N_s} - v) \pm \mathcal{A}(u - v, u_{N_s} - v)) \\
&\leqslant \mathcal{C}_1(a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min} \|u - v\|_{L^2_\rho(\Gamma;H^1_0(U))} \|u_{N_s} - v\|_{L^2_\rho(\Gamma;H^1_0(U))} \\
&+ |\mathcal{A}(v, u_{N_s} - v) - \mathcal{A}_{N_s}(v, u_{N_s} - v)|)
\end{aligned}$$

where $\mathcal{C}_1 := \frac{C_P(U)^2}{a_{min}\mathbb{F}^d_{min}\mathbb{F}^{-2}_{max}}$. Now, pick $v = u$, thus

$$\begin{aligned}
\|u - u_{N_s}\|_{L^2_\rho(\Gamma;H^1_0(U))} &\leqslant \mathcal{C}_1 \sup_{w \in L^2_\rho(\Gamma;H^1_0(U))} \frac{|\mathcal{A}(u, w) - \mathcal{A}_{N_s}(u, w)|}{\|w\|_{L^2_\rho(\Gamma;H^1_0(U))}} \\
&\leqslant \mathcal{C}_1 \sup_{x \in U, \mathbf{y} \in \Gamma} \|G(\mathbf{y}_s, \mathbf{y}_f) - G(\mathbf{y}_s)\| \|u\|_{L^2_\rho(\Gamma;H^1_0(U))}.
\end{aligned}$$

For notational simplicity we rewrite (7) as

$$\partial F(\mathbf{y}_s, \mathbf{y}_f) = I + A^s_{N_s}(\mathbf{y}_s) + A^f_{N_f}(\mathbf{y}_f)$$

for some set of matrices $A^s_{N_s}, A^f_{N_f} \in \mathbb{R}^{d\times d} \times U \times \Gamma$. With a slight abuse of notation we refer to $\partial F(\mathbf{y}_s)$ as $\partial F(\mathbf{y}_s) := I + A^s_{N_s}(\mathbf{y}_s)$. Note that $\mathbb{F}_{min} \leqslant \sigma_{min}(\partial F(\mathbf{y}_s, \mathbf{y}_f)) \Rightarrow \mathbb{F}_{min} \leqslant \sigma_{min}(\partial F(\mathbf{y}_s))$ and $\sigma_{min}(\partial F(\mathbf{y}_s, \mathbf{y}_f)) \leqslant \mathbb{F}_{max} \Rightarrow \sigma_{min}(\partial F(\mathbf{y}_s)) \leqslant \mathbb{F}_{max}$.

We now estimate the term $\|G(\mathbf{y}_s, \mathbf{y}_f) - G(\mathbf{y}_s)\|$. Denoting $J(\mathbf{y}_s, \mathbf{y}_f) := |\partial F(\mathbf{y}_s, \mathbf{y}_f)|\, \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1}$ $\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T}$ and similarly for $J(\mathbf{y}_s)$ we have

$$
\begin{aligned}
\|G(\mathbf{y}_s, \mathbf{y}_f) - G(\mathbf{y}_s)\| &= \|a(\mathbf{y}_s, \mathbf{y}_f) J(\mathbf{y}_s, \mathbf{y}_f) - a(\mathbf{y}_s) J(\mathbf{y}_s) \pm a(\mathbf{y}_s, \mathbf{y}_f) J(\mathbf{y}_s)]\| \\
&\leqslant |a(\mathbf{y}_s, \mathbf{y}_f) - a(\mathbf{y}_s)| \|J(\mathbf{y}_s)\| + a_{max} \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\| \\
&\leqslant \|a\|_{W^{1,\infty}(U)} |F(\mathbf{y}_s, \mathbf{y}_f) - F(\mathbf{y}_s)| + a_{max} \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\| \\
&\leqslant \|a\|_{W^{1,\infty}(U)} |\hat{v}| \sum_{i=N_s+1}^{N} \sqrt{\lambda_i} \|b_l(x)\|_{L^\infty(U)} \\
&\quad + a_{max} \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\|.
\end{aligned} \tag{32}
$$

Now,

$$
\begin{aligned}
\|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\| &\leqslant \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s) \pm |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \\
&\leqslant |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \|\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \\
&\quad + \big| |\partial F(\mathbf{y}_s, \mathbf{y}_f)| - |\partial F(\mathbf{y}_s)| \big| \|\partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \\
&\leqslant \mathbb{F}^d_{max} \|\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \\
&\quad + \mathbb{F}^{-2}_{min} \big| |\partial F(\mathbf{y}_s, \mathbf{y}_f)| - |\partial F(\mathbf{y}_s)| \big|.
\end{aligned} \tag{33}
$$

Applying the matrix identity $(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}$ where $A = I + A^s_{N_s}(\mathbf{y}_s)$, $B = -A^f_{N_f}(\mathbf{y}_f)$ and $C = D = I$ we obtain

$$
\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} = \partial F(\mathbf{y}_s)^{-1} + E(\mathbf{y}_s, \mathbf{y}_f)
$$

where

$$
\begin{aligned}
E(\mathbf{y}_s, \mathbf{y}_f) &:= -\partial F(\mathbf{y}_s)^{-1} A^f_{N_f}(\mathbf{y}_f) (I + \partial F(\mathbf{y}_s)^{-1} A^f_{N_f}(\mathbf{y}_f))^{-1} \partial F(\mathbf{y}_s)^{-1} \\
&= -\partial F(\mathbf{y}_s)^{-1} A^f_{N_f}(\mathbf{y}_f) \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1},
\end{aligned}
$$

then

$$
\begin{aligned}
\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T} &= E(\mathbf{y}_s, \mathbf{y}_f) E(\mathbf{y}_s, \mathbf{y}_f)^T \\
&\quad + \partial F(\mathbf{y}_s)^{-1} E(\mathbf{y}_s, \mathbf{y}_f)^T \\
&\quad + E(\mathbf{y}_s, \mathbf{y}_f) \partial F(\mathbf{y}_s)^{-T}.
\end{aligned}
$$

Now,

$$
\|E(\mathbf{y}_s, \mathbf{y}_f)\| \leqslant \mathbb{F}^{-2}_{min} \sum_{i=N_s+1}^{N} \sqrt{\lambda_i} \|B_i(x)\|
$$

It follows that

$$
\|\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \leqslant B_{\mathbb{T}} \mathbb{F}^{-3}_{min} (2 + \mathbb{F}^{-1}_{min}(1 - \tilde{\delta})) \tag{34}
$$

From Theorem 2.12 in [18] ($A, E \in \mathbb{C}^{d\times d}$ then $|det(A+E) - det(A)| \leqslant d\|E\| \max\{\|A\|, \|A+E\|\}^{d-1}$) we obtain

$$
\big| |\partial F(\mathbf{y}_s, \mathbf{y}_f)| - |\partial F(\mathbf{y}_s)| \big| \leqslant \mathbb{F}^{d-1}_{max} \mathbb{F}^{-2}_{min} B_{\mathbb{T}} d. \tag{35}
$$

Combining eqns (33), (34) and (35) we obtain

$$\|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\|_{L^\infty(U\times\Gamma)} \leqslant B_{\mathbb{T}} H(\mathbb{F}_{max}, \mathbb{F}_{min}, \tilde{\delta}, d).$$

The result follows. □

5.2. **Finite Element Error (II).** The second quantity controls the convergence with respect to the mesh size h. This will be determined by the polynomial order of the finite element subspace $H_h(U) \subset H_0^1(U)$ and the regularity of the solution u. From (23) we obtain the following bound:

$$\|u(\mathbf{y}_s) - u_h(\mathbf{y}_s)\|_{L^2_\rho(\Gamma_s;H_0^1(U))} \leqslant C_{\Gamma_s}(r)h^r$$

for some constant $r \in \mathbb{N}$ and $C_{\Gamma_s}(r) := \int_{\Gamma_s} C(r, u(\mathbf{y}_s))\rho(\mathbf{y}_s)d\mathbf{y}_s$. The constant r depends on the polynomial degree of the finite element basis and the regularity properties of the solution u (which is dependent on the regularity of f, the diffusion coefficient a and the mapping F). Similarly the error for the influence function is characterized as

$$\|\varphi(\mathbf{y}_s) - \varphi_h(\mathbf{y}_s)\|_{L^2_\rho(\Gamma_s;H_0^1(U))} \leqslant D_{\Gamma_s}(r)h^r$$

where $D_{\Gamma_s}(r) := \int_{\Gamma_s} C(r, \varphi(\mathbf{y}_s))\rho(\mathbf{y})d\mathbf{y}$. Following duality arguments and from Lemma 7 we obtain

$$E[|Q(u(\mathbf{y}_s)) - Q(u_h(\mathbf{y}_s))|] \leqslant a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min}C_{\Gamma_s}(r)D_{\Gamma_s}(r)h^{2r}. \tag{36}$$

5.3. **Sparse Grid Error (III).** In this section we shall not enumerate all the convergence rates that depend on the formulas from Table 1, but refer the reader to the appropriate citations. However, we will only explicitly derive the convergence rates for the isotropic Smolyak sparse grid.

Given the bounded linear functional Q we have that

$$\|Q(u_h(\mathbf{y}_s)) - Q(\mathcal{S}^{m,g}_w[u_h(\mathbf{y}_s)])\|_{L^2_\rho(\Gamma_s)} \leqslant a_{max}\mathbb{F}^d_{max}\mathbb{F}^{-2}_{min}\|e\|_{L^2_\rho(\Gamma_s;H_0^1(U))},$$

where $e := u_h(\mathbf{y}_s) - \mathcal{S}^{m,g}_w[u_h(\mathbf{y}_s)]$. However, as noted in Section 4.1, the sparse grid is computed with respect to the auxiliary density function $\hat{\rho}$, thus

$$\|e\|_{L^2_\rho(\Gamma_s;H_0^1(U))} \leqslant \left\|\frac{\rho}{\hat{\rho}}\right\|_{L^\infty(\Gamma_s)} \|e\|_{L^2_{\hat{\rho}}(\Gamma_s;H_0^1(U))}$$

The error term $\|e\|_{L^2_{\hat{\rho}}(\Gamma_s;H_0^1(U))}$ is controlled by the number of collocation knots η (or work), the choice of the approximation formulas $(m(i), g(\mathbf{i}))$ from Table 1, and the region of analyticity of $\Theta_\beta \subset \mathbb{C}^{N_s}$. From Theorem 1 the solution $u(\mathbf{y}_s)$ admits an extension in $\mathbb{C}^{N_s}$ i.e. $\mathbf{y}_s \to \mathbf{z}_s \in \mathbb{C}^{N_s}$ and $u(\mathbf{z}_s) \in C^0(\Theta_\beta; H_0^1(U))$. All the results proved in Section 3 can be obtained also for the semi-discrete solution $u_h(\mathbf{y}_s)$ which admits an analytic extension in the same region Θ_β and $u_h(\mathbf{z}_s) \in C^0(\Theta_\beta; H_h(U))$.

In [14, 19] the authors derive error estimates for isotropic and anisotropic Smolyak sparse grids with Clenshaw-Curtis and Gaussian abscissas where $\|e\|_{L^2_{\hat{\rho}}(\Gamma_s;H_0^1(U))}$ exhibit algebraic or sub-exponential convergence with respect to the number of collocation knots η (See Theorems 3.10, 3.11, 3.18 and 3.19 for more details). However, for these estimates to be valid the solution u has to admit and extension on a polyellipse in $\mathbb{C}^{N_s}$, $\mathcal{E}_{\sigma_1,\dots,\sigma_{N_s}} := \Pi_{i=1}^{N_s}\mathcal{E}_{n,\sigma_n}$, where

$$\mathcal{E}_{n,\sigma_n} = \left\{ z \in \mathbb{C};\ \mathrm{Re}(z) = \frac{e^{\sigma_n} + e^{-\sigma_n}}{2} cos(\theta),\ \ \mathrm{Im}(z) = \frac{e^{\sigma_n} - e^{-\sigma_n}}{2} sin(\theta), \theta \in [0, 2\pi) \right\},$$

and $\sigma_n > 0$. For an isotropic sparse grid the overall asymptotic subexponential decay rate $\hat{\sigma}$ will be dominated by the smallest σ_n i.e.

$$\hat{\sigma} \equiv \min_{n=1,\dots,N_s} \sigma_n.$$

Then the goal is to choose the largest $\hat{\sigma}$ such that $\mathcal{E}_{\sigma_1,\dots,\sigma_{N_s}} \subset \Theta_\beta$. First, recall from Section 3 that

$$\Theta_\beta := \left\{ \mathbf{z} \in \mathbb{C}^N;\, \mathbf{z} = \mathbf{y} + \mathbf{w},\, \mathbf{y} \in [-1,1]^{N_s},\, \sum_{l=1}^{N_s} \sup_{x \in \tilde{U}} \|B_l(x)\|_2 \sqrt{\lambda_l} |w_l| \leqslant \beta \} \right\}.$$

We can now form the set $\Sigma \subset \mathbb{C}^{N_s}$ such that $\Sigma \subset \Theta_\beta$, where $\Sigma := \Sigma_1 \times \cdots \times \Sigma_{N_s}$ and

$$\Sigma_n := \left\{ \mathbf{z} \in \mathbb{C};\, \mathbf{z} = \mathbf{y} + \mathbf{w},\, \mathbf{y} \in [-1,1],\, |w_n| \leqslant \tau_n := \frac{\beta}{1 - \tilde{\delta}} \right\}.$$

for $n = 1, \dots, N_s$. The polyellipse $\mathcal{E}_{\sigma_1,\dots,\sigma_n}$ can now be embedded in Σ by choosing $\sigma_1 = \sigma_2 = \cdots = \sigma_{N_s} = \hat{\sigma} = \log\left(\sqrt{\tau_{N_s}^2 + 1} + \tau_{N_s}\right) > 0$.

From Theorem 3.11 [19], given a sufficiently large η for a nested CC sparse grid we obtain the following estimate

$$\|e\|_{L^2_\rho(\Gamma_s; H^1_0(U))} \leqslant \mathcal{Q}(\sigma, \delta^*, N_s) \eta^{\mu_3(\sigma,\delta^*,N_s)} \exp\left(-\frac{N_s \sigma}{2^{1/N_s}} \eta^{\mu_2(N_s)} \right) \tag{37}$$

where

$$\mathcal{Q}(\sigma, \delta^*, N_s) := \frac{C_1(\sigma,\delta^*)}{\exp(\sigma\delta^* \tilde{C}_2(\sigma))} \frac{\max\{1, C_1(\sigma,\delta^*)\}^{N_s}}{|1 - C_1(\sigma,\delta^*)|},$$

$\sigma = \hat{\sigma}/2$, $\mu_2(N_s) = \frac{log(2)}{N_s(1+log(2N_s))}$ and $\mu_3(\sigma,\delta^*,N_s,) = \frac{\sigma\delta^*\tilde{C}_2(\sigma)}{1+\log(2N_s)}$. The constants $C_1(\sigma,\delta^*)$, $\tilde{C}_2(\sigma)$ and δ^* are defined in [19] eqns (3.11) and (3.12).

6. Complexity and Tolerance

In this section we derive the total work W needed such that $|var[Q(\mathbf{y}_s,\mathbf{y}_f)] - var[\mathcal{S}^{m,g}_w[Q_h(\mathbf{y}_s)]]|$ and $|\mathbb{E}[Q(\mathbf{y}_s,\mathbf{y}_f)] - \mathbb{E}[\mathcal{S}^{m,g}_w\ [Q_h(\mathbf{y}_s)]]|$ for the isotropic CC sparse grid is less or equal to a given tolerance parameter $tol \in \mathbb{R}^+$.

Let N_h be the number of degrees of freedom to solve the semi-discrete approximation $u_h \in H_h(U) \subset H^1_0(U)$. We assume that the complexity for solving for u_h is $\mathcal{O}(N_h^q)$, where the constant $q \geqslant 1$ reflects the optimality of the finite element solver. The cost of solving the approximation of the influence function $\varphi_h \in H_h(U)$ is also $\mathcal{O}(N_h^q)$. Thus for any $\mathbf{y}_s \in \Gamma_s$, the cost for computing $Q_h(\mathbf{y}_s) := B(\mathbf{y}_s; u_h(\mathbf{y}_s), \varphi_h(\mathbf{y}_s))$ is $\mathcal{O}(N_h^q)$.

Let $\mathcal{S}^{m,g}_w$ be the sparse grid operator characterized by $m(i)$ and $g(\mathbf{i})$. Furthermore, let $\eta(N_s, m, g, w, \Theta_\beta)$ be the number of the sparse grid knots. The total work for computing the variance $\mathbb{E}[(\mathcal{S}^{m,g}_w[Q_h(\mathbf{y}_s)])^2] - \mathbb{E}[\mathcal{S}^{m,g}_w[Q_h(\mathbf{y}_s)]]^2$ and the mean term $\mathbb{E}[\mathcal{S}^{m,g}_w[Q_h(\mathbf{y}_s)]]$ with respect to a given user tolerance is

$$W_{Total}(tol) = D_1 N_h^q(tol)\eta(tol)$$

for some constant $D_1 > 0$. We now separate the analysis into three parts:

(a) **Truncation:** From the truncation estimate derived in section 5.1 we seek $\|Q(\mathbf{y}_s,\ \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)} \leqslant \frac{tol}{3C_T}$ with respect to the decay of λ_i. First, make the assumption that $B_{\mathbb{T}} = \sup_{x\in U} \sum_{i=N_s+1}^{N} \sqrt{\lambda_i} \|B_i(x)\|_2 \leqslant C_D N_s^{-l}$ for some uniformly bounded $C_D > 0$. Furthermore, assume that $\|b_i(x)\|_{L^\infty(U)} \leqslant D_D \sup_{x\in U} \|B_i(x)\|_2$ for $i = 1, \dots N$ where $D_D > 0$ is uniformly bounded, thus $\sup_{x \in U} \sum_{i=N_s+1}^{N} \sqrt{\lambda_i} \|b_i(x)\|_2 \leqslant C_D D_D N_s^{-l}$. It follows that $\|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)} \leqslant \frac{tol}{3C_T}$ if

$$B_{\mathbb{T}} \leqslant C_D N_s^{-l} \leqslant D_2 tol$$

for some constant $D_2 > 0$. Finally, we have that

$$N_s(tol) \geqslant \left\lceil \left(\frac{D_2 tol}{C_D} \right)^{-1/l} \right\rceil.$$

(b) **Finite Element:** From Section 5.2 if

$$h(tol) \leqslant \left(\frac{tol}{3C_{FE} a_{min} \mathbb{F}^d_{min} \mathbb{F}^{-2}_{max} C_{\Gamma_s}(r) D_{\Gamma_s}(r)} \right)^{1/2r}$$

then $\|Q(\mathbf{y}_s) - Q_h(\mathbf{y}_s)\|_{L^2_\rho(\Gamma; H^1_0(U))} \leqslant \frac{tol}{3C_{FE}}$. Assuming that N_h grows as $\mathcal{O}(h^{-d})$ then

$$N_h(tol) \geqslant \left\lceil D_3 \left(\frac{tol}{3C_{FE} a_{min} \mathbb{F}^d_{min} \mathbb{F}^{-2}_{max} C_{\Gamma_s}(r) D_{\Gamma_s}(r)} \right)^{-d/2r} \right\rceil$$

for some constant $D_3 > 0$.

(c) **Sparse Grid:** Following the same strategy as in [19] (eqn (3.39)), to simplify the bound (37) choose $\delta^* = (e \log(2) - 1)/\tilde{C}_2(\sigma)$ and $\tilde{C}_2(\sigma)$. Thus $\|e\|_{L^2_{\hat{\rho}}(\Gamma_s; H^1_0(U))} \leqslant \frac{tol}{3C_{SG} C_T} \left\| \frac{\rho}{\hat{\rho}} \right\|^{-1}_{L^\infty(\Gamma_s)}$ if

$$\eta(tol) \geqslant \left\lceil \left(\frac{3\|\rho/\hat{\rho}\|_{L^\infty(\Gamma_s)} C_{SG} C_T C_F F^{N_s} \exp(\sigma(\beta, \tilde{\delta}))}{tol} \right)^{\frac{1+\log(2N_s)}{\sigma}} \right\rceil$$

where $C_F = \frac{C_1(\sigma,\delta^*)}{|1 - C_1(\sigma,\delta^*)|}$ and $F = \max\{1, C_1(\sigma, \delta^*)\}$.

Combining (a), (b) and (c) we obtain that for a given user error tolerance *tol* the total work is

$$\begin{aligned} W_{Total}(tol) &= D_1 N_h^q(tol, D_3) \eta(\tilde{\delta}, \beta, N_s(tol), \|\rho/\hat{\rho}\|_{L^\infty(\Gamma_s)}) \\ &= \mathcal{O}\left(\left(\frac{\|\rho/\hat{\rho}\|_{L^\infty(\Gamma_s)} F^{C tol^{-1/l}}}{tol} \right)^{\sigma^{-1}(1 + \log 2C - l^{-1} \log tol)} \right). \end{aligned}$$

for some $C > 0$.

7. Numerical Results

We test our method on a square domain. Suppose the reference domain is set $U = (0,1) \times (0,1)$ and stochastically deforms according to the following rule:

$$\begin{array}{ll} F(x_1, x_2) = (x_1, \, (x_2 - 0.5)(1 + ce(\omega, x_1)) + 0.5) & if \quad x_2 > 0.5 \\ F(x_1, x_2) = (x_1, \, x_2) & if \quad 0 \leqslant x_2 \leqslant 0.5 \end{array}$$

for some positive constant $c > 0$. In other words we deform only the upper half of the domain and fix the button half. We set the Dirichlet boundary conditions to zero everywhere except at the upper border to $u(x_1, x_2)|_{\partial D(\omega)} = g(x_1)$, where $g(x_1) := exp(\frac{-1}{(1-(2(x_1-0.5))^2})$ (See Figure 2). This implies that the value at the upper boundary does not change with boundary perturbation but the solution does become stochastic with respect to the domain perturbation. Consider a QoI defined on the bottom half of the reference domain, which is not deformed, as

$$Q(u) := \int_{(0,1)} \int_{(0,1/2)} g(x_1) g(2x_2) u(\omega, x_1, x_2)\, dx_1 dx_2.$$

We now show a numerical example with linear decay on the gradient of the deformation, i.e. the gradient terms $\sqrt{\lambda_n} \sup_{x \in U} \|B_n(x)\|$ decay linearly as n^{-1}.

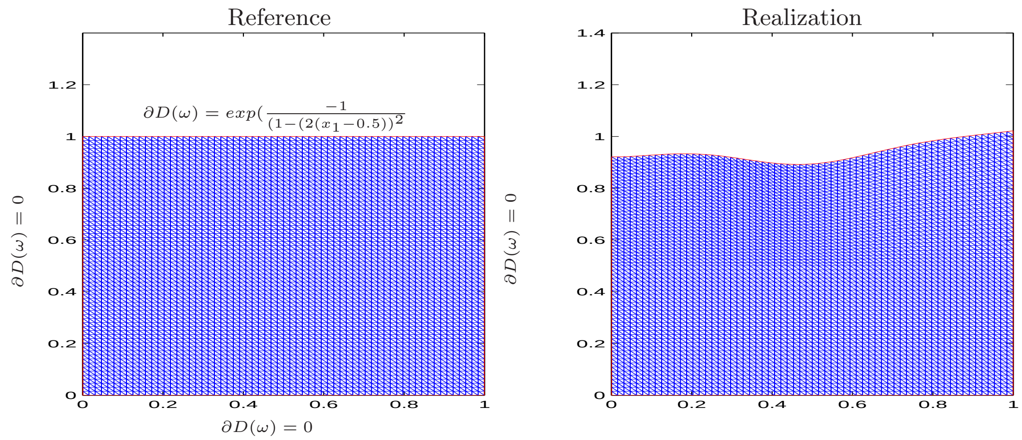


FIGURE 2. Stochastic deformation of a square domain. (left) Reference square domain with Dirichlet boundary conditions. (right) Vertical deformation from stochastic model.

(a) $a(x) = 1$ for all $x \in U$.
(b) Stochastic Model

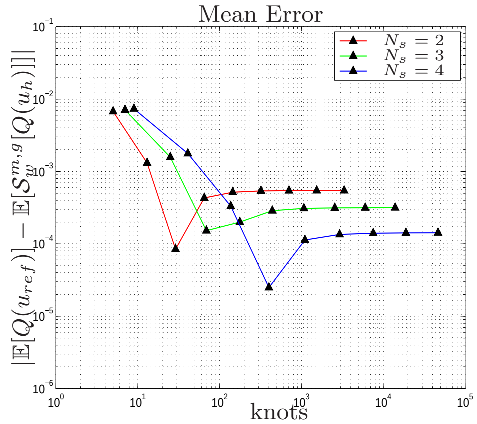
$$e_S(\omega, x_1) := Y_1(\omega) \left(\frac{\sqrt{\pi} L}{2} \right) + \sum_{n=2}^{N_s} \sqrt{\lambda_n} \varphi_i(x_1) Y_n(\omega);$$
$$e_F(\omega, x_1) := \sum_{n=N_s+1}^{N} \sqrt{\lambda_n} \varphi_n(x_1) Y_n(\omega)$$

(c) Linear decay $\sqrt{\lambda_n} := \frac{(\sqrt{\pi} L)^{1/2}}{n}$, $n \in \mathbb{N}$.

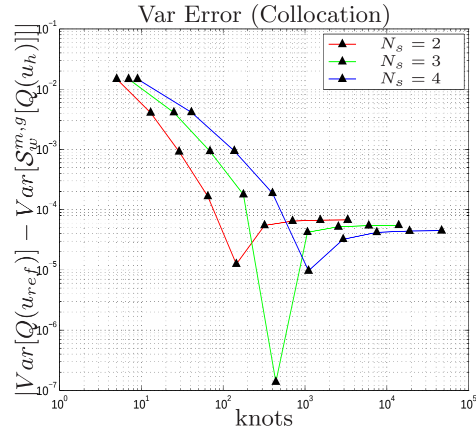
$$\varphi_i(x_1) := \begin{cases} n^{-1} sin\left(\frac{\lfloor n/2 \rfloor \pi x_1}{L_p}\right) & \text{if n is even} \\ n^{-1} cos\left(\frac{\lfloor n/2 \rfloor \pi x_1}{L_p}\right) & \text{if n is odd} \end{cases}$$

This implies that $\sup_{x \in U} \sigma_{max}(B_l(x))$ is bounded by a constant and we obtain linear decay on the gradient of the deformation.

(d) $\{Y_n\}_{n=1}^N$ are independent uniform distributed in $(-\sqrt{3}, \sqrt{3})$
(e) $L = 19/50$, $L_P = 1$, $c = 0.1533$, $N = 15$.
(f) 129×129 triangular mesh
(g) $\mathbb{E}[Q_h]$ and $\mathrm{var}(Q_h)$, are computed with a Clenshaw-Curtis isotropic sparse grid (Sparse Grid Toolbox V5.1, [20, 21]).
(h) The reference solutions $\mathrm{var}[Q_h(u_{ref})]$ and $\mathbb{E}[Q_h(u_{ref})]$ are computed with an adaptive Sparse Grid ($\approx 30,000$ knots) [22] with a 257×257 mesh for $N = 15$ dimensions.
(i) The QoI is normalized by the reference solution $Q(U)$.



(a)



(b)

FIGURE 3. Collocation results for $N_s = 2, 3, 4$ with linear decay. (a) Mean error with respect to reference. Observe that the convergence rate decays exponentially until truncation saturation is reached. (b) Variance error for with respect to reference. For this case we also observe that the convergence rate is faster than polynomial.

In Figure 3 we show the results of the matlab code for $N_s = 2, 3, 4$ and compare the results with respect to a $N = 15$ dimensional adaptive sparse grid method collocation with $\approx 30,000$ collocation points [22]. The computed mean value is 1.0152 and variance is 0.0293 (0.17 std).

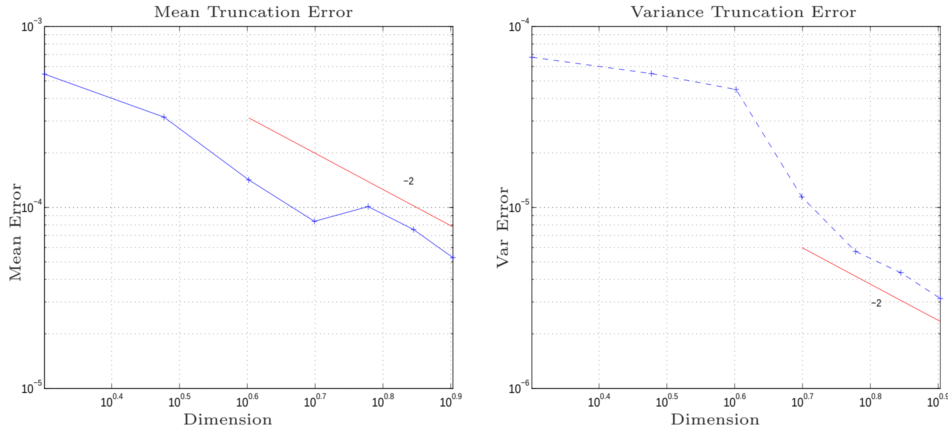
In Figure 3 (a) and (b) the normalized mean and variance errors are shown for $N_s = 2, 3, 4$. For (a) notice the exponential decay from the sparse grid approximation until the truncation error and/or finite element error starts to dominate. In (b) the variance error decay is actually subexponential as indicated by the error bounded in (37).

We now analyze the decay of the truncation error. For $N_s = 2, 3, 4, 5$ we compute the mean and variance error as in (g). However, for $N_s = 6, 7, 8$ we increase the mesh size to 257x257 vertices and we choose an adaptive sparse grid with 15,000 sparse grid points to compute the mean and variance. This should be enough to make the error contribution from the finite element and sparse grid error very small compared to the truncation error. The reference solution for the mean and variance is computed as in part (h).

In Figure 4 we plot the truncation error for (a) the mean and (b) the variance with respect to the number of dimensions. We observe that we obtain a convergence rate that appears faster than the linear decay of the gradient of the stochastic deformation. This indicates we can further improve the convergence rate of the truncation estimate.

8. CONCLUSIONS

In this paper we give a rigorous convergence analysis of the stochastic collocation approach based on isotropic Smolyak grids for the approximation of an elliptic PDE defined on a random domain. This consists of an analysis of the regularity of the solution with respect to the parameters describing the domain perturbation. Moreover, we derive error estimates both in the "energy norm"



(a) (b)

FIGURE 4. Truncation Error with respect to the number of dimension. (a) Mean error. (b) Variance error. In both cases the decay appears faster than linear, which is faster than the predicted convergence rate.

as well as on functionals of the solution (Quantity of Interest) for Clenshaw Curtis abscissas that can be easily generalized to a larger class of sparse grids.

We show that for a linear elliptic partial differential equation with a random domain the solution can be analytically extended to a well defined region Θ_β embedded in $\mathbb{C}^N$ with respect to the random variables. This analysis leads to a provable subexponential convergence rate of the QoI computed with an isotropic Clenshaw-Curtis sparse grid. We show that the size of this region, and the rate of convergence, is directly related to the decay of the gradient of the stochastic deformation.

As our numerical experiments demonstrate, we are able to solve the mean and variance of the QoI with moderate deformations of the domain (leading to a coefficient of variation of the QoI of ≈ 0.17). This is a clear advantage over the perturbation approaches that are restricted to small deviations. In addition, the numerical experiments confirm the sub-exponential rate predicted from the error estimates.

This approach is well suited for a moderate number of stochastic variables but becomes impractical for large problems. However, we can easily extend this approach to anisotropic sparse grids [14].

Acknowledgements

We appreciate the advice from Quan Long and Serge Prudhomme.

References

[1] Z. Zhenhai, J. White, A fast stochastic integral equation solver for modeling the rough surface effect computer-aided design, in: IEEE/ACM International Conference ICCAD-2005, 2005, pp. 675–682.
[2] J. S. H. C. Chauviere, L. Lurati., Computational modeling of uncertainty in time-domain electromagnetics, SIAM J. Sci. Comput. 28 (2006) 751–775.
[3] D. Fransos, Stochastic numerical methods for wind engineering, Ph.D. thesis, Politecnico di Torino (2008).

[4] D. Tartakovsky, D. Xiu, Stochastic analysis of transport in tubes with rough walls, Journal of Computational Physics 217 (1) (2006) 248 – 259, uncertainty Quantification in Simulation Science.
[5] H. Harbrecht, R. Schneider, C. Schwab, Sparse second moment analysis for elliptic problems in stochastic domains, Numerische Mathematik 109 (2008) 385–414.
[6] L. C. Evans, Partial Differential Equations, Vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, Rhode Island, 1998.
[7] I. C. F. Ipsen, D. J. Lee, Determinant approximations, ArXiv (arXiv:1105.0437v1).
[8] D. London, A note on matrices with positive definite real part, Proceedings of the American Mathematical Society 82 (3) (1981) pp. 322–324.
[9] S. G. Krantz, Function Theory of Several Complex Variables, AMS Chelsea Publishing, Providence, Rhode Island, 1992.
[10] R. Gunning, H. Rossi, Analytic Functions of Several Complex Variables, American Mathematical Society, 1965.
[11] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Mathematics, Doklady 4 (1963) 240–243.
[12] V. Barthelmann, E. Novak, K. Ritter, High dimensional polynomial interpolation on sparse grids, Advances in Computational Mathematics 12 (2000) 273–288.
[13] J. Bäck, F. Nobile, L. Tamellini, R. Tempone, Stochastic spectral galerkin and collocation methods for PDEs with random coefficients: A numerical comparison, in: J. S. Hesthaven, E. M. Rnquist (Eds.), Spectral and High Order Methods for Partial Differential Equations, Vol. 76 of Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2011, pp. 43–62.
[14] F. Nobile, R. Tempone, C. Webster, An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data, SIAM Journal on Numerical Analysis 46 (5) (2008) 2411–2442. arXiv:http://epubs.siam.org/doi/pdf/10.1137/070680540.
[15] J. Charrier, Strong and weak error estimates for elliptic partial differential equations with random coefficients, SIAM J. Numer. Anal. 50 (1) (2012) 216–246.
[16] I. Babuska, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM Journal on Numerical Analysis 42 (2) (2004) 800–825. arXiv:http://epubs.siam.org/doi/pdf/10.1137/S0036142902418680.
[17] P. Frauenfelder, C. Schwab, R. A. Todor, Finite elements for elliptic problems with stochastic coefficients, Computer Methods in Applied Mechanics and Engineering 194 (2-5) (2005) 205 – 228, selected papers from the 11th Conference on The Mathematics of Finite Elements and Applications.
[18] I. Ipsen, R. Rehman, Perturbation bounds for determinants and characteristic polynomials, SIAM Journal on Matrix Analysis and Applications 30 (2) (2008) 762–776. arXiv:http://epubs.siam.org/doi/pdf/10.1137/070704770.
[19] F. Nobile, R. Tempone, C. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM Journal on Numerical Analysis 46 (5) (2008) 2309–2345. arXiv:http://epubs.siam.org/doi/pdf/10.1137/060663660.
[20] A. Klimke, B. Wohlmuth, Algorithm 847: spinterp: Piecewise multilinear hierarchical sparse grid interpolation in MATLAB, ACM Transactions on Mathematical Software 31 (4).
[21] A. Klimke, Sparse Grid Interpolation Toolbox – user's guide, Tech. Rep. IANS report 2007/017, University of Stuttgart (2007).
[22] T. Gerstner, M. Griebel, Dimension-adaptive tensor-product quadrature, Computing 71 (1) (2003) 65–87.

Recent publications :

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

33.2013 P. CHEN, A. QUARTERONI, G. ROZZA:
Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations

34.2013 A. ABDULLE, M. J. GROTE, C. STOHRER:
Finite element heterogeneous multiscale method for the wave equation: long time effects

35.2013 A. CHKIFA, A. COHEN, G. MIGLIORATI, F. NOBILE, R. TEMPONE:
Discreate least squares polynomial approximation with random evaluations-application to parametric and stochastic elliptic PDEs

36.2013 N. GUGLIELMI, D. KRESSNER, C. LUBICH:
Computing extremal points of symplectic pseudospectra and solving symplectic matrix nearness problems

37.2013 S. DEPARIS, D. FORTI, A. QUARTERONI:
A rescaled localized radial basis functions interpolation on non-cartesian and non-conforming grids

38.2013 A. GIZZI, R. RUIZ-BAIER, S. ROSSI, A. LAADHARI, C. CHERUBINI, S. FILIPPI:
A three-dimensional continuum model of active contraction in single cardiomyocytes

39.2013 P. TRICERRI, L. DEDÈ, A. QUARTERONI, A. SEQUEIRA:
Numerical validation of isotropic and transversely isotropic constitutive models for healthy and unhealthy cerebral arterial tissues

40.2013 D. KRESSNER, M. STEINLECHNER, A. USCHMAJEW:
Low-rank tensor methods with subspace correction for symmetric eigenvalue problems

41.2013 A. ABDULLE, O. BUDÁC:
An adaptive finite element heterogeneous multiscale method for Stokes flow in porous media

42.2013 C. JÄGGLI, L.IAPICHINO, G. ROZZA:
An improvement on geometrical parametrizations by transfinite maps

43.2013 P. PACCIARINI, G. ROZZA:
Stabilized reduced basis method for parametrized advection-diffusion PDEs

44.2013 A. KOSHAKJI, A. QUARTERONI, G. ROZZA:
Free form deformation techniques applied to 3D shape optimization problems

45.2013 J. E. CATRILLON-CANDAS, F. NOBILE, R. F. TEMPONE:
Analytic regularity and collocation approximation for PDEs with random domain deformations