

# Analytic regularity and collocation approximation for elliptic PDEs with random domain deformations

Julio E. Castrillón-Candás<sup>a,\*</sup>, Fabio Nobile<sup>b</sup>, Raúl F. Tempone<sup>c,a</sup>

<sup>a</sup>*SRI Center for Uncertainty Quantification, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.*

<sup>b</sup>*CSQI-MATHICSE, Ecole Polytechnique Federale Lausanne, Station 8, CH1015, Lausanne, Switzerland*

<sup>c</sup>*Applied Mathematics and Computational Science, 4700, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia*

---

## 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.

*Keywords:* Uncertainty Quantification, Stochastic Collocation, Stochastic PDEs, Finite Elements, Complex Analysis, Smolyak Sparse Grids

---

## 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 graphene 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 [1–5]. The collocation approaches proposed in [2–4] work well for large amplitude domain perturbations

---

\*corresponding author

*Email addresses:* `uvel@alum.mit.edu` (Julio E. Castrillón-Candás), `fabio.nobile@epfl.ch` (Fabio Nobile), `raul.tempone@kaust.edu.sa` (Raúl F. Tempone)

*Preprint submitted to Computers and Mathematics with applications*

*November 16, 2018*

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 [1, 5] are efficient for small domains perturbation.

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. In Section 3 we show that the solution can be analytically extended into a well defined region in  $\mathbb{C}^N$ . Theorem 7 is the main result of this paper. In Section 4 we setup the stochastic collocation problem and summarize several known sparse grid 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  $N_s \leq N$  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  $\Omega$  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  $L_P^q(\Omega)$ ,  $q \in [1, \infty]$ , as the space of random variables such that

$$L_P^q(\Omega) := \left\{ v \mid \int_{\Omega} |v(\omega)|^q d\mathbb{P} < \infty \right\} \text{ and } L_P^\infty(\Omega) := \{v \mid \text{ess sup}_{\omega \in \Omega} |v(\omega)| < \infty\},$$

where  $v : \Omega \rightarrow \mathbb{R}$  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 sufficiently smooth regularity on  $f(\cdot, \omega), a(\cdot, \omega) : D(\omega) \rightarrow \mathbb{R}^d$ , find  $u(\cdot, \omega) : D(\omega) \rightarrow \mathbb{R}$  such that almost surely

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

Now, assume the diffusion coefficient satisfies the following assumption.

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

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

where

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

We now state the weak formulation as:

**Problem 1.** Find  $u(\cdot, \omega) \in H_0^1(D(\omega))$  s.t.

$$\int_{D(\omega)} a(x, \omega) \nabla u(x, \omega) \cdot \nabla v(x) \, dx = \int_{D(\omega)} f(x, \omega) v(x) \, dx \quad \forall v \in H_0^1(D(\omega)) \text{ a.s. in } \Omega, \quad (1)$$

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  $\Omega$ .

### 2.1. Reformulation onto a fixed Domain

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

**Assumption 2.** Suppose that the map  $F(\omega) : U \rightarrow D(\omega)$  is one-to-one a.s. and that there exist constants  $\mathbb{F}_{min}$  and  $\mathbb{F}_{max}$  such that

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

almost everywhere in  $U$  and almost surely in  $\Omega$ . 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  $\Omega$  and a.e. in  $U$  unless disambiguation is needed. Moreover, for any random function  $v(\omega) : D(\omega) \rightarrow \mathbb{R}$ , we denote by  $\hat{v} = v \circ F$  the random function  $\hat{v}(\cdot, \omega) = v(F(\cdot, \omega), \omega) : U \rightarrow \mathbb{R}$ .

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

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

**Remark 1.** Note that we refer to  $\nabla v : D(\omega) \rightarrow \mathbb{R}^d$  as the gradient of  $v$  on  $D(\omega)$  and  $\nabla \hat{v} : U \rightarrow \mathbb{R}^d$ , with  $\hat{v} = v \circ F$ , as the gradient of  $\hat{v}$  on  $U$ . Therefore the notation  $\nabla(v \circ F) : U \rightarrow \mathbb{R}^d$  refers to the gradient of  $\hat{v}$  on  $U$ .

Now let us recall the chain rule for Sobolev spaces (see Theorem 3.35 in [6]): Let  $U, D \subset \mathbb{R}^n$  and  $T : U \rightarrow D$  be invertible. Suppose that  $T$  and  $T^{-1}$  have continuous, bounded derivatives of order less or equal to  $j$ . Then if  $k \in W^{j,p}(D)$ ,  $p \geq 1$ , we have  $v = k \circ T \in W^{j,p}(U)$  and the derivatives of  $v$  are given by the chain rule.

Thus by the chain rule on Sobolev spaces for any  $v \in H^1(D(\omega))$  we have that  $\nabla v = \partial F^{-T} \nabla(v \circ F)$ , where  $v \circ F \in H^1(U)$ . By a change of variables, the weak form can now be posed as:

**Problem 2.** Find  $\hat{u}(\cdot, \omega) \in H_0^1(U)$  s.t.

$$B(\omega; \hat{u}(\cdot, \omega), v) = l(\omega; v), \quad \forall v \in H_0^1(U), \quad (3)$$

where for any  $v, s \in H_0^1(U)$

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

$$l(\omega; v) := \int_U (f \circ F)(\cdot, \omega) v |\partial F(\cdot, \omega)|,$$

$(f \circ F)(\cdot, \omega) \in L^2(U)$  and  $C(\cdot, \omega) = \partial F^T(\cdot, \omega) \partial F(\cdot, \omega)$  almost surely. We now recover  $u(\cdot, \omega) : D(\omega) \rightarrow H_0^1(\omega)$  as  $u = \hat{u} \circ F^{-1}$ .

Note that under this notation  $\hat{u}(\cdot, \omega)$  can be written as  $u(\cdot, \omega) \circ F(\cdot, \omega)$  or shortly as  $u \circ F$ , which is the notation used in the rest of the paper. Note that we can also use the notation  $(u \circ F)(\cdot, \omega)$ .

The following lemmas give the conditions under which Problem 2 is well posed.

**Lemma 1.** Under Assumptions 2 we have that

- i)  $L^2(D(\omega))$  and  $L^2(U)$  are isomorphic.
- ii)  $H^1(D(\omega))$  and  $H^1(U)$  are isomorphic.

Moreover,  $\forall v \in H^1(D(\omega))$

$$\|\nabla v\|_{L^2(D(\omega))} \leq \mathbb{F}_{max}^{d/2} \mathbb{F}_{min}^{-1} \|\nabla(v \circ F)\|_{L^2(U)}. \quad (4)$$

PROOF. i) is immediate. Now, from 2 and the chain rule on Sobolev spaces we obtain that  $\forall v \in H_0^1(D(\omega))$  the inequality (4) is satisfied. We can similarly obtain a bound for the converse. It follows that  $H^1(D(\omega))$  and  $H^1(U)$  are isomorphic.

**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

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

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

PROOF. From Assumption 2 we have that

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

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

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

and

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

Thus Problem 2 is uniformly continuous and coercive. Furthermore, since  $f \circ F \in L^2(U)$  then 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  $H_0^1(U)$  and  $H_0^1(D(\omega))$  (Lemma 1).

From the Cauchy-Schwartz inequality we obtain

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

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

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

Now, from equation (4) it follows that  $\forall v \in H_0^1(D(\omega))$

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

**Remark 2.** 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 we have the following boundary value problem: Given  $f(\cdot, \omega), a(\cdot, \omega) : D(\omega) \rightarrow \mathbb{R}^d$  and  $g(\cdot, \omega) : \partial D(\omega) \rightarrow \mathbb{R}^d$  find  $u(\cdot, \omega) : D(\omega) \rightarrow \mathbb{R}^d$  such that almost surely

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

Since  $U$  is bounded and Lipschitz there exists a bounded linear operator  $T : H^{1/2}(\partial U) \rightarrow H^1(U)$  such that  $\forall \tilde{g} \in H^{1/2}(\partial U)$  we have that  $\tilde{\mathbf{w}} := T\tilde{g} \in H^1(U)$  satisfies  $\tilde{\mathbf{w}}|_U = \tilde{g}$  almost surely. The weak formulation can now be posed as ([7] chapter 6, p297):

**Problem 3.** Given that  $f \circ F \in L^2(U)$  find  $\hat{u}(\cdot, \omega) \in H_0^1(U)$  s.t.

$$B(\omega; \hat{u}, v) = \tilde{l}(\omega; v), \quad \forall v \in H_0^1(U)$$

almost surely, where  $\tilde{l}(\omega; v) := \int_U (f \circ F)(\cdot, \omega) |\partial F(\cdot, \omega)| v - L(\hat{\mathbf{w}}(\cdot, \omega), v)$ ,  $\hat{g} := g \circ F$ ,  $\hat{\mathbf{w}} := T(\hat{g})$ ,

$$L(\hat{\mathbf{w}}(\cdot, \omega), v) := \int_U (a \circ F)(\cdot, \omega) (\nabla(\hat{\mathbf{w}}(\cdot, \omega)))^T C^{-1}(\cdot, \omega) |\partial F(\cdot, \omega)| \nabla v,$$

and  $\hat{\mathbf{w}}(\cdot, \omega)|_{\partial U} = \hat{g}(\cdot, \omega)$ . This homogeneous boundary value problem can be remapped to  $D(\omega)$  as  $\tilde{u}(\cdot, \omega) := (\hat{u} \circ F^{-1})(\cdot, \omega)$ , thus we can rewrite  $\hat{u}(\cdot, \omega) = (\tilde{u} \circ F)(\cdot, \omega)$ .

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

### 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

$$Q(u) := \int_{\tilde{D}} q(x)u(x, \omega) dx \quad (5)$$

with  $q \in L^2(\tilde{D})$  over the region  $\tilde{D} \subset D(\omega)$  for any  $\omega \in \Omega$ . Moreover, we assume that  $\exists \delta > 0$  such that  $\text{dist}(\tilde{D}, D(\omega)) < \delta \forall \omega \in \Omega$  and  $F|_{\tilde{D}} = I$  on  $\tilde{D}$  (i.e. there is no deformation of the domain on  $\tilde{D}$ ).

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

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

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

$$B(\omega; v, \varphi) = Q(v) \quad (6)$$

a.s. in  $\Omega$ .

We can now pick a particular  $T$  such that  $\hat{\mathbf{w}} = T(\hat{g})$  vanishes inside  $\tilde{D}$ . Therefore, we have that

$$Q(u) = Q(\tilde{u} + \hat{\mathbf{w}} \circ F^{-1}) = Q(\tilde{u}),$$

and thus

$$Q(\tilde{u}) = \int_{\tilde{D}} q(x)\tilde{u}(x, \omega) dx = \int_{\tilde{D}} q \circ F(x)\tilde{u} \circ F(x, \omega) dx = Q(\tilde{u} \circ F) = B(\omega; \tilde{u} \circ F, \varphi).$$

### 2.2. Domain Parameterization

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

Define the induced measure  $\mu_Y$  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  $\Gamma$ , then there exists a density function  $\rho(\mathbf{y}) : \Gamma \rightarrow [0, +\infty)$  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

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

Further, we define the following spaces:

$$L_\rho^q(\Gamma) := \left\{ v \mid \int_{\Omega} |v(\mathbf{y})|^q \rho(\mathbf{y}) d\mathbf{y} < \infty \right\} \text{ and } L_\rho^\infty(\Gamma) := \left\{ v \mid \text{ess sup}_{\mathbf{y} \in \Gamma} |v(\mathbf{y})| < \infty \right\}.$$

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

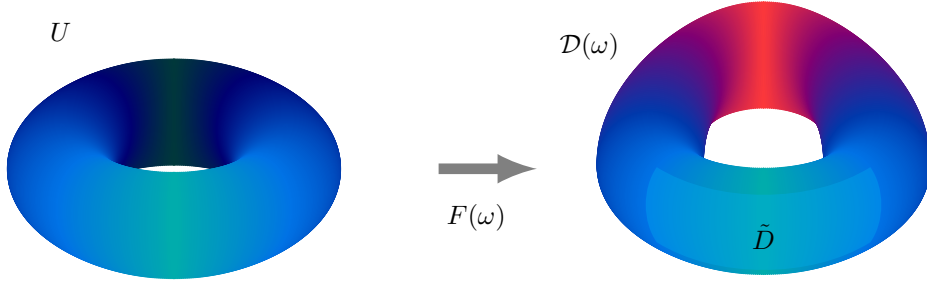


Figure 1: Cartoon example of stochastic domain realization from a reference domain. Note that the non stochastic area  $\tilde{D}$  is contained in the interior of  $D(\omega)$ . Note that this figure is modified from the TikZ tex code from *Smooth map of manifolds and smooth spaces* by Andrew Stacey.

**Assumption 3.** The map  $F(\omega) : U \rightarrow D(\omega)$  has the form

$$F(x, \omega) := x + e(x, \omega)\hat{v}(x)$$

a.s. in  $\Omega$ , with  $\hat{v} : U \rightarrow \mathbb{R}^d$ ,  $\hat{v} := [\hat{v}_1, \dots, \hat{v}_d]^T$ ,  $\hat{v}_i \in C^1(U)$  for  $i = 1, \dots, d$ , and  $e(\cdot, \omega) : U \rightarrow D(\omega)$ . Assume that the map  $F(\omega) : U \rightarrow D(\omega)$  is one-to-one almost surely.

We next assume the stochastic perturbation  $e(x, \omega)$  with the following parameterization:

$$e(x, \omega) := \sum_{l=1}^N \sqrt{\mu_l} b_l(x) Y_l(\omega). \quad (7)$$

Denote  $Y := [Y_1, \dots, Y_N]$ , and for  $n = 1, \dots, N$  let  $\Gamma_n \equiv Y_n(\Omega)$  and  $\mathbb{E}[Y_n] = 0$ . Furthermore denote  $\Gamma := \prod_{n=1}^N \Gamma_n$ , and  $\rho(\mathbf{y}) : \Gamma \rightarrow \mathbb{R}_+$  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)  $b_1, \dots, b_N \in W^{1, \infty}(U)$ .
- 3)  $\|b_l\|_{L^\infty(U)} = 1$  for  $l = 1, 2, \dots, N$ .
- 4)  $\mu_l$  are monotonically decreasing for  $l = 1, 2, \dots, N$ .

Note that the parameterization of equation (7) may arise from a truncated Karhunen Loeve (KL) expansion of mean square random fields. However, in general KL expansion eigenfunctions  $b_l$  belong to  $L^2(U)$ , but can be extended to  $L^\infty(U)$  or even higher regularity if the covariance function satisfies certain regularity properties (see [8] details).

### 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. In this section we analyze the analyticity of the solution assuming a stochastic domain but with deterministic  $f$ .

From the stochastic model formulated in Section 2 the Jacobian  $\partial F$  is written as

$$\partial F(x, \omega) = I + \sum_{l=1}^N B_l(x) \sqrt{\mu_l} Y_l(\omega) \quad (8)$$

with

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

where  $\partial \hat{v}$  is the Jacobian of  $\hat{v}(x)$ .

#### Assumption 5.

a)  $a \circ F$  and  $\hat{g}$  are only a function of  $x \in U$  and independent of  $\omega \in \Omega$ .

b) There exists  $0 < \tilde{\delta} < 1$  such that  $\sum_{l=1}^N \|B_l(x)\|_2 \sqrt{\mu_l} \leq 1 - \tilde{\delta}$ ,  $\forall x \in U$ .

**Remark 3.** Assumption 5 a) restricts  $a(x, \omega)$  (and  $\hat{g}(x, \omega)$ ) to be a constant along the direction  $\hat{v}(x)$ . This assumption simplifies the presentation of this section. However, Assumption 5 a) is still useful for many practical problems. For example, in layered materials the width or geometry of each layer is stochastic, but the diffusion coefficient for each layer is known. This example can be found in semi-conductor design [1]. Moreover, this assumption allows the diffusion coefficient to be non constant along the non stochastic directions.

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

$$\Theta_\beta := \left\{ \mathbf{z} \in \mathbb{C}^N; \mathbf{z} = \mathbf{y} + \mathbf{w}, \mathbf{y} \in [-1, 1]^N, \sum_{l=1}^N \sup_{x \in U} \|B_l(x)\|_2 \sqrt{\mu_l} |w_l| \leq \beta \right\}, \quad (9)$$

where  $\mathbf{w} := [w_1, \dots, w_N] \in \mathbb{C}^N$ .

**Assumption 6.** Assume that  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  can be analytically extended in  $\mathbb{C}^d$ . Furthermore assume that the analytic extension  $\text{Re}(f \circ F)(\cdot, \mathbf{z}), \text{Im}(f \circ F)(\cdot, \mathbf{z}) \in L^2(U) \forall \mathbf{z} \in \Theta_\beta$ .



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  $R(\mathbf{z})$  and similarly for  $F(x, \mathbf{y})$  unless emphasis is needed. We shall now prove several lemmas that will be useful to prove the main results (Theorem 7).

Note that we apply many singular value inequalities and properties to prove Lemmas 5, 6 and parts of section 5, that can be found in [9]. We recall some of the most commonly used. If  $A, B \in \mathbb{C}^{n,n}$  then

1.  $\sigma_{max}(A + B) \leq \sigma_{max}(A) + \sigma_{max}(B)$ .
2.  $\sigma_{min}(A + B) \geq \sigma_{min}(A) - \sigma_{max}(B)$ .
3.  $\sigma_{max}(AB) \leq \sigma_{max}(A)\sigma_{max}(B)$ .
4.  $\sigma_{min}(AB) \geq \sigma_{min}(A)\sigma_{min}(B)$ .

In addition, if  $A, B$  are Hermitian then

1.  $\lambda_{max}(A + B) \leq \lambda_{max}(A) + \lambda_{max}(B)$  (Lidskii inequality).
2.  $\lambda_{min}(A + B) \geq \lambda_{min}(A) + \lambda_{min}(B)$  (Dual Lidskii inequality).

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

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

PROOF. i)  $\|\partial F(\mathbf{y})\|_2 \leq 1 + \sup_{x \in U} \sum_{l=1}^N \|B_l(x)\|_2 \sqrt{\mu_l} \leq 2 - \tilde{\delta}$ .

ii)  $\sigma_{max}(\partial F(\mathbf{y}) - I) = \|\sum_{l=1}^N B_l(x) \sqrt{\mu_l} y_l\|_2 \leq 1 - \tilde{\delta} \Rightarrow \sigma_{min}(\partial F(\mathbf{y})) = \sigma_{min}(I + \partial F(\mathbf{y}) - I) \geq \sigma_{min}(I) - \sigma_{max}(\partial F(\mathbf{y}) - I) \geq 1 - (1 - \tilde{\delta}) = \tilde{\delta}$ .

iii) The result follows from the following fact: If  $A \in \mathbb{C}^{d \times d}$  we have that  $\sigma_{min}(A) \leq |\lambda_l(A)| \leq \sigma_{max}(A)$  for all  $l = 1, \dots, d$  (see [9]).

**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  $\forall x \in U$  we have that  $\det(\partial F(\mathbf{z}))$  is analytic and*

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

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

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

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 [10]: If  $A \in \mathbb{C}^{d \times d}$  and  $\sigma_{max}(A) < 1$  then

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

It follows that

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

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

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

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

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  $\det(\partial F(\mathbf{z}))$  is a finite polynomial of  $\mathbf{z}$ . The rest of the result follows by applying Lemma 3.

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

$$0 < \beta < \min \left\{ \frac{\tilde{\delta} \log(2 - \gamma)}{d + \log(2 - \gamma)}, \sqrt{1 + \tilde{\delta}^2/2} - 1 \right\}, \quad (10)$$

where  $\gamma := \frac{2\tilde{\delta}^2 + (2 - \tilde{\delta})^d}{\tilde{\delta}^d + (2 - \tilde{\delta})^d}$  then  $\forall x \in U$  we have that  $\text{Re } G(\mathbf{z})$  is positive definite  $\forall \mathbf{z} \in \Theta_\beta$  and

(a)  $\lambda_{\min}(\text{Re } G(\mathbf{z})^{-1}) \geq B(\tilde{\delta}, \beta, d, a_{\max}) > 0$ , where

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

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

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

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

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

**PROOF.** (a) To simplify the proof we use the property that if  $\text{Re } G^{-1}(\mathbf{z})$  is positive definite then  $\text{Re } G(\mathbf{z})$  is positive definite (From (b) in [11]), but first we derive bounds for  $\text{Re } \partial F(\mathbf{z})^T \partial F(\mathbf{z})$  and  $\text{Im } \partial F(\mathbf{z})^T \partial F(\mathbf{z})$ . Now,  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$  we have that

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

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  $\lambda_{\min}(A + B) \geq \lambda_{\min}(A) + \lambda_{\min}(B)$ ) we obtain

$$\begin{aligned} \lambda_{\min}(\text{Re } \partial F(\mathbf{z})^T \partial F(\mathbf{z})) &\geq \lambda_{\min}((I + R(\mathbf{y}) + R_r(\mathbf{w}))^T (I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\ &\quad - \lambda_{\max}(R_i(\mathbf{w})^T R_i(\mathbf{w})) \\ &= \sigma_{\min}^2(I + R(\mathbf{y}) + R_r(\mathbf{w})) - \sigma_{\max}^2(R_i(\mathbf{w})) \\ &\geq (\sigma_{\min}(I + R(\mathbf{y})) - \sigma_{\max}(R_r(\mathbf{w})))^2 - \sigma_{\max}^2(R_i(\mathbf{w})) \\ &\geq (\tilde{\delta} - \beta)^2 - \beta^2. \end{aligned} \quad (11)$$

It follows that if  $\beta < \tilde{\delta}/2$  then  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

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

and is positive definite. We see that  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$ ,

$$\begin{aligned} \max_{i=1,\dots,d} |\lambda_i(\text{Im } \partial F(\mathbf{z})^T \partial F(\mathbf{z}))| &\leq \sigma_{\max}(R_i(\mathbf{w})^T(I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\ &\quad + (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w}) \\ &\leq 2\sigma_{\max}(R_i(\mathbf{w}))\sigma_{\max}(I + R(\mathbf{y}) + R_r(\mathbf{w})) \\ &\leq 2\beta(2 + (\beta - \tilde{\delta})). \end{aligned} \quad (12)$$

We now have that  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

$$\begin{aligned} \lambda_{\min}((a \circ F)^{-1} \text{Re}(\xi^{-1}(\mathbf{z}) \partial F(\mathbf{z})^T \partial F(\mathbf{z}))) &\geq \frac{1}{a_{\max} |\xi(\mathbf{z})|^2} \lambda_{\min}(\xi_R(\mathbf{z}) \text{Re } F(\mathbf{z})^T \partial F(\mathbf{z})) \\ &\quad + \xi_I(\mathbf{z}) \text{Im } F(\mathbf{z})^T \partial F(\mathbf{z}) \\ &\geq \frac{1}{a_{\max} |\xi(\mathbf{z})|^2} (\xi_R(\mathbf{z}) \lambda_{\min}(\text{Re } \partial F(\mathbf{z})^T \partial F(\mathbf{z})) \\ &\quad - |\xi_I(\mathbf{z})| \max_{i=1,\dots,d} \lambda_i(\text{Im } \partial F(\mathbf{z})^T \partial F(\mathbf{z}))), \end{aligned} \quad (13)$$

where  $\xi(\mathbf{z}) := \xi_R(\mathbf{z}) + i\xi_I(\mathbf{z}) = \det(I + R(\mathbf{z}))$ . From Lemma 5  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$  we have that  $|\xi(\mathbf{z})|^{-1} \geq (2 - \tilde{\delta})^{-d}(2 - \alpha)^{-1} > 0$  whenever  $\mathbf{z} \in \Theta_\beta$  and thus

$$\begin{aligned} \max_{i=1,\dots,d} |\lambda_i(\text{Im } \partial F(\mathbf{z})^T \partial F(\mathbf{z}))| &\leq \max_{i=1,\dots,d} |\lambda_i(R_i(\mathbf{w})^T(I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\ &\quad + (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w}))| \\ &\leq 2\beta(2 + (\beta - \tilde{\delta})). \end{aligned} \quad (14)$$

From Lemma 3 (iii) if  $\beta < \frac{\tilde{\delta} \log \gamma}{d + \log \gamma}$ ,  $\gamma := \frac{2\tilde{\delta}^d + (2 - \tilde{\delta})^d}{\tilde{\delta}^d + (2 - \tilde{\delta})^d}$ , then  $\xi_R(\mathbf{z}) > |\xi_I(\mathbf{z})|$ ,  $\forall \mathbf{z} \in \Theta_\beta$  and  $\forall \mathbf{z} \in \Theta_\beta$ . From inequalities (11) and (12) we have that if  $\beta < \sqrt{1 + \tilde{\delta}^2}/2 - 1$  then  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

$$\lambda_{\min}(\text{Re } \partial F(\mathbf{z})^T \partial F(\mathbf{z})) > \max_{i=1,\dots,d} |\lambda_i(\text{Im } \partial F(\mathbf{z})^T \partial F(\mathbf{z}))|$$

and  $\lambda_{\min}(\text{Re } G(\mathbf{z})^{-1}) \geq B(\tilde{\delta}, \beta, d, a_{\max}) > 0$ , where

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

From London's Lemma [11] it follows that  $\text{Re } G(\mathbf{z})$  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  $\lambda_{\max}(A + B) \leq \lambda_{\max}(A) + \lambda_{\max}(B)$ ) we have that  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

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

From inequalities (15) and (14), and Lemmas 3 and 4, we obtain  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

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

where

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

(c) Similarly,  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$ , we can bound

$$\begin{aligned} \sigma_{max}(\operatorname{Im} \partial F(\mathbf{z})^T \partial F(\mathbf{z})) &\leq \sigma_{max}(R_i(\mathbf{w})^T (I + R(\mathbf{y}) + R_r(\mathbf{w})) \\ &\quad + (I + R(\mathbf{y}) + R_r(\mathbf{w}))^T R_i(\mathbf{w})) \\ &\leq 2\sigma_{max}(R_i(\mathbf{w})) \sigma_{max}(I + R(\mathbf{y}) + R_r(\mathbf{w})) \\ &\leq 2\beta(2 + (\beta - \tilde{\delta})). \end{aligned} \tag{16}$$

and

$$\begin{aligned} \sigma_{max}(\operatorname{Re} \partial F(\mathbf{z})^T \partial F(\mathbf{z})) &\leq \sigma_{max}((I + R(\mathbf{y}) + R_r(\mathbf{w}))^T (I + R(\mathbf{y}) + R_r(\mathbf{w}))) \\ &\quad + \sigma_{max}(R_i(\mathbf{w})^T R_i(\mathbf{w})) \\ &= \sigma_{max}^2(I + R(\mathbf{y}) + R_r(\mathbf{w})) + \sigma_{max}^2(R_i(\mathbf{w})) \\ &\leq (\sigma_{max}(I + R(\mathbf{y})) + \sigma_{max}(R_r(\mathbf{w})))^2 + \sigma_{max}^2(R_i(\mathbf{w})) \\ &\leq ((2 - \tilde{\delta}) + \beta)^2 + \beta^2. \end{aligned} \tag{17}$$

From inequalities (16) and (17), and Lemmas 3 and 4 we obtain  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

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

where

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

**Lemma 6.** For all  $\mathbf{z} \in \Theta_\beta$  and  $x \in U$

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

where

$$\varepsilon(\tilde{\delta}, \beta, d, a_{max}, a_{min}) := \frac{1}{\left(1 + \left(\frac{C(\tilde{\delta}, \beta, d, a_{min})}{B(\tilde{\delta}, \beta, d, a_{max})}\right)^2\right)} D(\tilde{\delta}, \beta, d, a_{min}). \quad (18)$$

PROOF. From Lemma 5  $\text{Re } G(\mathbf{z})$  is positive definite  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$ , where  $\beta$  satisfies (10). It follows from the Lemma in [11] that  $G(\mathbf{z}) = Q(I + i\Lambda)Q^*$ , where  $Q$  is a non-singular matrix,  $\Lambda := \text{diag}(\alpha_1, \dots, \alpha_d)$  and  $\alpha_1, \dots, \alpha_d$  are real. Since  $G(\mathbf{z})$  is symmetric then  $\text{Re } G(\mathbf{z}) = (1/2)(G(\mathbf{z}) + G(\mathbf{z})^*)$  and it is simple to see that  $\text{Re } G(\mathbf{z}) = QQ^*$ . Note that in [11], the real part of a generic complex matrix  $A$  (i.e.  $(1/2)(A + A^*)$ ) is not in general  $\text{Re } A$ .

We need now to show that  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$   $\lambda_{min}(\text{Re } G(\mathbf{z})) = \sigma_{min}^2(Q) \geq \varepsilon > 0$ , with  $\varepsilon$  defined in equation (18). Applying (b) in [11] we have that

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

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

$$\lambda_{max}(\text{Re } G(\mathbf{z})^{-1}) = \sigma_{max}^2(DQ^{-1}) \geq \sigma_{min}^2(D)\sigma_{max}^2(Q^{-1}) = \sigma_{min}^2(D)\sigma_{min}^{-2}(Q),$$

and therefore

$$\sigma_{min}^2(Q) \geq \frac{\sigma_{min}^2(D)}{\lambda_{max}(\text{Re } G(\mathbf{z})^{-1})} = \frac{\sigma_{min}((I + \Lambda^2)^{-1})}{\lambda_{max}(\text{Re } G(\mathbf{z})^{-1})} \geq \frac{(1 + \sigma_{max}^2(\Lambda))^{-1}}{\lambda_{max}(\text{Re } G(\mathbf{z})^{-1})}.$$

Now,  $\text{Im } G(\mathbf{z})^{-1} = (DQ^{-1})^*(-\Lambda)DQ^{-1}$  and  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

$$\sigma_{max}(\text{Im } G(\mathbf{z})^{-1}) \geq \sigma_{min}^2(DQ^{-1})\sigma_{max}(\Lambda).$$

Since  $\text{Re } G(\mathbf{z})^{-1} = (DQ^{-1})^*DQ^{-1}$  then  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$   $\lambda_{min}(\text{Re } G(\mathbf{z})^{-1}) = \sigma_{min}^2(DQ^{-1})$  and

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

It follows that  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

$$\lambda_{min}(\text{Re } G(\mathbf{z})) \geq \frac{1}{(1 + A^2)|\lambda_{max}(\text{Re } G(\mathbf{z})^{-1})|}. \quad (19)$$

From Lemma 5 (a) we have that  $\lambda_{min}(\text{Re } G(\mathbf{z})^{-1}) \geq B(\tilde{\delta}, \beta, d, a_{max}) > 0$ . From Lemma 5 (c) we have that  $\sigma_{max}(\text{Im } G(\mathbf{z})^{-1}) \leq C(\tilde{\delta}, \beta, d, a_{min}) < \infty$ . This implies  $\sigma_{max}(\Lambda) \leq A(\tilde{\delta}, \beta, d, a_{max}, a_{min}) < \infty$ . Finally from Lemma 5 (b)  $\lambda_{max}(\text{Re } G(\mathbf{z})^{-1}) \leq D(\tilde{\delta}, \beta, d, a_{min}) < \infty$ . We conclude that  $\forall x \in U$  and  $\forall \mathbf{z} \in \Theta_\beta$

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

We are now ready to prove the main result of this section.

**Theorem 7.** Let  $0 < \tilde{\delta} < 1$  then the solution  $\tilde{u} \circ F : \Gamma \rightarrow H_0^1(U)$  of Problem 3 can be extended holomorphically in  $\Theta_\beta \subset \mathbb{C}^N$  (see equation (9)) if

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

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

PROOF. The strategy for this proof is to show that the function  $(\tilde{u} \circ F)(\cdot, \mathbf{y})$  can be extended on  $\Theta_\beta$  and is analytic in each variable separately. Then apply Hartog's Theorem (Chap1, p32, [12]) and Osgood's Lemma (Chap 1, p 2, [13]) to show that such an extension is analytic in  $\Theta_\beta$ .

For  $n = 1, \dots, N$  consider the map  $\Psi(\mathbf{y}) : \Gamma \rightarrow H_0^1(U)$  where

$$\Psi(\mathbf{y}) := (\tilde{u} \circ F)(\mathbf{y}),$$

for any arbitrary points  $\mathbf{y} \in \Gamma$ . Furthermore, consider the extension of  $\mathbf{y} \rightarrow \mathbf{z}$ , where  $\mathbf{z} \in \mathbb{C}^N$ .

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 in  $\Theta_\beta$ . From Lemma 4 it follows that the entries of  $G(\mathbf{z})$  are analytic on  $\Theta_\beta$ .

Let  $\bar{\Psi}(\mathbf{z}) := [\bar{\Psi}_R(\mathbf{z}), \bar{\Psi}_I(\mathbf{z})]^T$ , with  $\bar{\Psi}_R(\mathbf{z}) := \text{Re } \Psi(\mathbf{z})$  and  $\bar{\Psi}_I(\mathbf{z}) := \text{Im } \Psi(\mathbf{z})$ , be the solution (in the weak sense) of the problem

$$-\nabla \cdot \hat{G}(\mathbf{z}) \nabla \bar{\Psi}(\mathbf{z}) = \hat{f}(\mathbf{z}), \quad (20)$$

where

$$\hat{G}(\mathbf{z}) := \begin{pmatrix} G_R(\mathbf{z}) & -G_I(\mathbf{z}) \\ G_I(\mathbf{z}) & G_R(\mathbf{z}) \end{pmatrix}, \quad \hat{f} := \begin{pmatrix} \tilde{f}_R(\mathbf{z}) \\ \tilde{f}_I(\mathbf{z}) \end{pmatrix},$$

$G_R(\mathbf{z}) := \text{Re}(G(\mathbf{z}))$ ,  $G_I(\mathbf{z}) := \text{Im}(G(\mathbf{z}))$ ,  $\tilde{f}_R(\mathbf{z}) := \text{Re } \tilde{f}(\mathbf{z})$  and  $\tilde{f}_I = \text{Im } \tilde{f}(\mathbf{z})$ . Note that  $\tilde{f}(\mathbf{z})$  refers to extension of the right hand side of the weak formulation i.e.  $\tilde{l}(\mathbf{z}; v)$  for all  $v \in H_0^1(U)$ . Thus  $\tilde{f}_R(\mathbf{z}) := \text{Re}\{(f \circ F)(\cdot, \mathbf{z})|\partial F(\mathbf{z})| + \nabla \cdot (a \circ F)C^{-1}(\mathbf{z})|\partial F(\mathbf{z})|\nabla \hat{\mathbf{w}}\}$  and similarly for  $\tilde{f}_I(\mathbf{z})$ .

The system of equations (20) has a unique solution. Indeed, from Lemma 6 we have that  $\lambda_{\min}(G_R(\mathbf{z})) > 0 \forall \mathbf{z} \in \Theta_\beta \forall x \in U$ . Since  $G_I(\mathbf{z})$  is symmetric, it follows that  $G(\mathbf{z})$  is positive definite  $\forall x \in U$ , hence the well posedness of problem (20) by Lax-Milgram. Moreover, the solution  $\Psi(\mathbf{z})$  coincides with  $(\tilde{u} \circ F)(\mathbf{y})$ ,  $\mathbf{y} \in \Gamma$ , when  $\mathbf{z} \in \Gamma$ , therefore it is a complex continuation of  $\tilde{u} \circ F$  on  $\Theta_\beta$ .

To show that  $\Psi(\mathbf{z}) : \Theta_\beta \rightarrow H_0^1(U)$  is holomorphic on  $\Theta_\beta$  we focus on the  $n^{\text{th}}$  variable  $z_n$ ,  $n = 1, \dots, N$ , which we write as  $z_n = s + iw$ ,  $s, w \in \mathbb{R}$ , and 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 (20) with respect to  $s = \operatorname{Re} z_n$  and  $w = \operatorname{Im} z_n$  we obtain

$$\begin{aligned}
-(\nabla \cdot G_R \nabla \partial_s \Psi_R(\mathbf{z}) - \nabla \cdot G_I \nabla \partial_s \Psi_I(\mathbf{z})) &= \nabla \cdot \partial_s G_R \nabla \Psi_R(\mathbf{z}) - \nabla \cdot \partial_s G_I \nabla \Psi_I(\mathbf{z}) \\
&\quad + \partial_s \tilde{f}_R(\mathbf{z}), \\
-(\nabla \cdot G_I \nabla \partial_s \Psi_R(\mathbf{z}) + \nabla \cdot G_R \nabla \partial_s \Psi_I(\mathbf{z})) &= \nabla \cdot \partial_s G_I \nabla \Psi_R(\mathbf{z}) + \nabla \cdot \partial_s G_R \nabla \Psi_I(\mathbf{z}) \\
&\quad + \partial_s \tilde{f}_I(\mathbf{z}), \\
-(\nabla \cdot G_R \nabla \partial_w \Psi_R(\mathbf{z}) - \nabla \cdot G_I \nabla \partial_w \Psi_I(\mathbf{z})) &= \nabla \cdot \partial_w G_R \nabla \Psi_R(\mathbf{z}) - \nabla \cdot \partial_w G_I \nabla \Psi_I(\mathbf{z}) \\
&\quad + \partial_w \tilde{f}_R(\mathbf{z}), \\
-(\nabla \cdot G_I \nabla \partial_w \Psi_R(\mathbf{z}) + \nabla \cdot G_R \nabla \partial_w \Psi_I(\mathbf{z})) &= \nabla \cdot \partial_w G_I \nabla \Psi_R(\mathbf{z}) + \nabla \cdot \partial_w G_R \nabla \Psi_I(\mathbf{z}) \\
&\quad + \partial_w \tilde{f}_I(\mathbf{z}). \tag{21}
\end{aligned}$$

Note that to avoid clutter in the equations we refer to  $G_R(\mathbf{z})$  as  $G_R$  and  $G_I(\mathbf{z})$  as  $G_I$ . By the Lax-Milgram theorem the derivatives  $\partial_s \Psi(\mathbf{z})$  and  $\partial_w \Psi(\mathbf{z})$  exist and have a unique solution whenever  $\mathbf{z} \in \Theta_\beta$ . The second step is now to show that the Cauchy-Riemann conditions are satisfied.

Let  $P(\mathbf{z}) := \partial_s \Psi_R(\mathbf{z}) - \partial_w \Psi_I(\mathbf{z})$  and  $Q(\mathbf{z}) := \partial_w \Psi_R(\mathbf{z}) + \partial_s \Psi_I(\mathbf{z})$ . To show analyticity we have to show that  $P(\mathbf{z}) = 0$  and  $Q(\mathbf{z}) = 0$  for all  $\mathbf{z} \in \Theta_\beta$ . By taking linear combinations of equations (21) we obtain

$$\begin{aligned}
-\nabla \cdot (G_R \nabla P - G_I \nabla Q) &= \nabla \cdot ((\partial_s G_R - \partial_w G_I) \nabla \Psi_R - (\partial_w G_R + \partial_s G_I) \nabla \Psi_I) \\
&\quad + \partial_s \tilde{f}_R - \partial_w \tilde{f}_I, \\
-\nabla \cdot (G_I \nabla P + G_R \nabla Q) &= \nabla \cdot ((\partial_w G_R + \partial_s G_I) \nabla \Psi_R - (\partial_s G_R - \partial_w G_I) \nabla \Psi_I) \\
&\quad + \partial_s \tilde{f}_I + \partial_w \tilde{f}_R. \tag{22}
\end{aligned}$$

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

From Assumption 5 we have that  $(f \circ F)(\mathbf{z})$  is analytic on  $\Theta_\beta$  thus  $\tilde{l}(\mathbf{z}; v)$  is holomorphic on  $\Theta_\beta$ . Now, recall that  $G(\mathbf{z})$  is analytic if  $\mathbf{z} \in \Theta_\beta$ . Thus equations (22) have a unique solution  $P(\mathbf{z}) = Q(\mathbf{z}) = 0$  for all  $\mathbf{z} \in \Theta_\beta$ .

From Hartog's Theorem it follows that  $\Psi(\mathbf{z})$  is continuous on  $\Theta_\beta$ . From Osgood's Lemma it follows that  $\Psi(\mathbf{z})$  is holomorphic on  $\Theta_\beta$ .

**Corollary 8.** *The following estimate holds for all  $\mathbf{z} \in \Theta_\beta$ :*

$$\|\nabla((\tilde{u} \circ F)(\cdot, \mathbf{z}))\|_{L^2(U)} \leq \frac{E((f \circ F)(\mathbf{z}), \hat{\mathbf{w}}, a_{max}, \tilde{\delta}, \alpha, d, C_P(U))}{\varepsilon(\tilde{\delta}, \beta, d, a_{max}, a_{min})}, \tag{23}$$

where

$$\begin{aligned}
E((f \circ F)(\mathbf{z}), \hat{\mathbf{w}}, a_{max}, \tilde{\delta}, \alpha, d, C_P(U)) &:= (2 - \alpha) \left( \frac{\|\nabla \hat{\mathbf{w}}\|_{L^2(U)}}{a_{max}^{-1} (2 - \tilde{\delta})^{-d} \tilde{\delta}^2} \right. \\
&\quad \left. + C_P(U) \|(f \circ F)(\mathbf{z})\|_{L^2(U)} \right)
\end{aligned}$$

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



PROOF. We formally multiply (20) by  $\bar{\Psi}(\mathbf{z})^T$  and integrate over  $U$  to obtain

$$\int_U (\nabla \Psi_R(\mathbf{z}))^T G_R(\mathbf{z}) \nabla \Psi_R(\mathbf{z}) + (\nabla \Psi_I(\mathbf{z}))^T G_R(\mathbf{z}) \nabla \Psi_I(\mathbf{z}) = \int_U \Psi_R(\mathbf{z}) \tilde{f}_R(\mathbf{z}) + \Psi_I(\mathbf{z}) \tilde{f}_I(\mathbf{z}).$$

From Lemma 3, Lemma 4 and the Poincaré inequality we have that

$$\begin{aligned} \left| \int_U \Psi_R(\mathbf{z}) \tilde{f}_R(\mathbf{z}) + \Psi_I(\mathbf{z}) \tilde{f}_I(\mathbf{z}) \right| &\leq \left| \int_U \operatorname{Re}\{\Psi(\mathbf{z})^* \tilde{f}(\mathbf{z})\} \right| \leq \int_U |(\nabla \hat{\mathbf{w}})^T G(\mathbf{z}) \nabla \Psi(\mathbf{z})^*| \\ &\quad + \int_U |\Psi(\mathbf{z})^* (f \circ F)(\mathbf{z})| \partial F(\mathbf{z})| \\ &\leq \frac{\|\nabla \hat{\mathbf{w}}\|_{L^2(U)} \|\nabla \Psi(\mathbf{z})\|_{L^2(U)}}{a_{max}^{-1} (2 - \tilde{\delta})^{-d} (2 - \alpha)^{-1} \tilde{\delta}^2} \\ &\quad + (2 - \tilde{\delta}) C_P(U) \|\nabla \Psi(\mathbf{z})\|_{L^2(U)} \|(f \circ F)(\mathbf{z})\|_{L^2(U)}. \end{aligned}$$

It follows that

$$\|\nabla \Psi(\mathbf{z})\|_{L^2(U)} \min_{x \in U} \{\lambda_{min}(G_R(\mathbf{z}))\} \leq E((f \circ F)(\mathbf{z}), \hat{\mathbf{w}}, a_{max}, \tilde{\delta}, \alpha, d).$$

From Lemma 6 the result follows.

**Remark 4.** We can relax the restrictions on  $f$  from Assumption 6 for Theorem 7. For example, suppose that the  $N_{\mathbf{q}}$  valued random vector  $\mathbf{q}$  takes values on  $\Gamma_{\mathbf{q}} := \tilde{\Gamma}_1 \times \dots \times \tilde{\Gamma}_{N_{\mathbf{q}}}$  (with the probability density  $\tilde{\rho}(\mathbf{q})$ ). Assume that the random vector  $\mathbf{q}$  is independent from  $\mathbf{y}$  and write  $f$  as

$$f(\cdot, \mathbf{q}) = \sum_{j=1}^{N_f} \tilde{y}_j(\mathbf{q}) \tilde{b}_j(\cdot),$$

where for  $j = 1, \dots, N_f$ ,  $\tilde{y}_j \in L_{\tilde{\rho}}^{\infty}(\Gamma_{\mathbf{q}})$  and  $\tilde{b}_j : \mathbb{R}^d \rightarrow \mathbb{R}$ . We now have stochastic contributions from the coefficients  $\tilde{y}_j(\mathbf{q})$ . Since  $\tilde{b}_j$  is defined on  $\mathbb{R}^d$  we can remap onto the reference domain and obtain

$$(f \circ F)(\cdot, \mathbf{y}, \mathbf{q}) = \sum_{j=1}^{N_f} \tilde{y}_j(\mathbf{q}) (\tilde{b}_j \circ F)(\cdot, \mathbf{y}).$$

Now, assume that for  $j = 1, \dots, N_f$ ,  $\tilde{b}_j$  can be analytically extended in  $\mathbb{C}^d$ . Furthermore, for  $j = 1, \dots, N_f$  the extensions  $\operatorname{Re}(\tilde{b}_j \circ F)(\cdot, \mathbf{z}), \operatorname{Im}(\tilde{b}_j \circ F)(\cdot, \mathbf{z}) \in L^2(U) \forall \mathbf{z} \in \Theta_{\beta}$ . This implies that we can holomorphically extend  $(\tilde{b}_j \circ F)(\mathbf{y})$  into  $\Theta_{\beta}$ . Now, suppose that the coefficients  $\tilde{y}_j(\mathbf{q})$  can be analytically extended into  $\mathbb{C}^{N_{\mathbf{q}}}$ . It follows that  $(f \circ F)(\cdot, \mathbf{y}, \mathbf{q})$  can be analytically extended in  $\Theta_{\beta} \times \mathbb{C}^{N_{\mathbf{q}}}$ .

Using a similar proof strategy (and linearity of the elliptic operator) as in Theorem 7, we can show that  $(\tilde{u} \circ F)(\cdot, \mathbf{y}, \mathbf{q})$  can be analytically extended along each separate dimension in  $\Theta_{\beta} \times \mathbb{C}^{N_{\mathbf{q}}}$ . By using Hartog's Theorem and Osgood's Lemma we can conclude that  $(\tilde{u} \circ F)(\cdot, \mathbf{y}, \mathbf{q})$  can be analytically extended in  $\Theta_{\beta} \times \mathbb{C}^{N_{\mathbf{q}}}$ .

#### 4. Stochastic Collocation

We seek to efficiently approximate the mean and variance of the QoI of the form (5). 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:

- $H_h(U) \subset H_0^1(U)$  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  $h > 0$ .
- $\mathcal{P}_p(\Gamma) \subset L_\rho^2(\Gamma)$  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  $N_p = \prod_{n=1}^N (p_n + 1)$ .

- $u_h : \Gamma \rightarrow H_h(U)$  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.,

$$\int_U (a \circ F)(\cdot, \mathbf{y}) [\nabla u_h(\mathbf{y})]^T C^{-1}(\mathbf{y}) \nabla v_h |\partial F|(\mathbf{y}) dx = \int_U (f \circ F)(\cdot, \mathbf{y}) v_h |\partial F|(\mathbf{y}) dx - L(\hat{\mathbf{w}}, v_h) \quad (24)$$

$\forall v_h \in H_h(U)$  and for a.e.  $\mathbf{y} \in \Gamma$ . Denote  $\pi_h : H_0^1(U) \rightarrow H_h(U)$  as the finite element operator s.t. if  $u \in H_0^1(U)$  then  $u_h := \pi_h u$  and

$$\|u - \pi_h u\|_{H_0^1(U)} \leq C_\pi \min_{v \in H_h(U)} \|u - v\|_{H_0^1(U)} \leq h^r C(r, u). \quad (25)$$

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.,

$$\int_U (a \circ F)(\cdot, \mathbf{y}) [\nabla v_h(\mathbf{y})]^T C^{-1}(\mathbf{y}) \nabla \varphi_h dx = Q(v_h) \quad \forall v_h \in H_h(U). \quad (26)$$

**Remark 5.** Note that for the sake of simplicity we ignore quadrature errors and assume that the integrals (24) and (26) are computed exactly.

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

$$\hat{\rho}(\mathbf{y}) = \prod_{n=1}^N \hat{\rho}_n(y_n) \quad \forall \mathbf{y} \in \Gamma, \quad \text{and is such that} \quad \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma)} < \infty. \quad (27)$$

For each dimension  $n = 1, \dots, N$ , let  $y_{n,k_n}$ ,  $1 \leq k_n \leq p_n + 1$ , 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  $\int_{\Gamma_n} q_{p_n+1}(\mathbf{y})v(\mathbf{y})\hat{\rho}_n(\mathbf{y})d\mathbf{y} = 0$  for all  $v \in \mathcal{P}_{p_n}(\Gamma_n)$ .

Standard choices for  $\hat{\rho}$ , such as constant, Gaussian, etc., lead to the 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  $[k_1, \dots, k_N]$  we associate the global index

$$k = k_1 + p_1(k_2 - 1) + p_1p_2(k_3 - 1) + \dots$$

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

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

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  $\mathcal{I}_p : C^0(\Gamma) \rightarrow \mathcal{P}_p(\Gamma)$ , such that

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

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

$$\mathbb{E}_{\hat{\rho}}^p[g] = \sum_{k=1}^{N_p} \omega_k g(y_k), \quad \omega_k = \prod_{n=1}^N \omega_{k_n}, \quad \omega_{k_n} = \int_{\Gamma_n} l_{k_n}^2(y)\hat{\rho}_n(y)dy. \quad (28)$$

In the case  $\rho/\hat{\rho}$  is a smooth function we can use directly (28) to approximate the mean value or the variance of  $Q_h$  as

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

Otherwise,  $\mathbb{E}[Q_h]$  and  $\text{var}_h(Q_h)$  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. [14, 15]) to build a multivariate polynomial approximation on a sparse grid. See [16] for details.

Let  $\mathcal{I}_n^{m(i)} : C^0(\Gamma_n) \rightarrow \mathcal{P}_{m(i)-1}(\Gamma_n)$  be the 1D interpolant as previously introduced. Here  $i \geq 1$  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  $i \geq 1$ . In addition, let  $m(0) = 0$  and  $\mathcal{I}_n^{m(0)} = 0$ . Further, we introduce the difference operators

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

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

$$\mathcal{S}_w^{m,g}[Q_h] = \sum_{\mathbf{i} \in \mathbb{N}_+^N : g(\mathbf{i}) \leq w} \bigotimes_{n=1}^N \Delta_n^{m(i_n)}(Q_h) \quad (29)$$

or equivalently written as

$$\mathcal{S}_w^{m,g}[Q_h] = \sum_{\mathbf{i} \in \mathbb{N}_+^N : g(\mathbf{i}) \leq w} c(\mathbf{i}) \bigotimes_{n=1}^N \mathcal{I}_n^{m(i_n)}(Q_h), \quad \text{with } c(\mathbf{i}) = \sum_{\substack{\mathbf{j} \in \{0,1\}^N \\ g(\mathbf{i}+\mathbf{j}) \leq w}} (-1)^{|\mathbf{j}|}. \quad (30)$$

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  $g(\mathbf{i}) \leq w$  in (30) is typically chosen so as to forbid the use of tensor grids of high degree in all directions at the same time.

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

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

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.

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

The following result proved in [16], 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

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

We recall that the most typical choice of  $m$  and  $g$  is given by (see [14, 15])

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

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 [16]).

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

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 [17] combines the advantages of isotropic sparse collocation with those of anisotropic full tensor product collocation. Note that in [18], the authors show convergence of sparse grid approximations for an elliptic PDE with random coefficients with infinite dimensions i.e.  $N = \infty$ .

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

$$\mathbb{E}[\mathcal{S}_w^{m,g} Q_h] = \mathbb{E}_{\hat{\rho}}[\mathcal{S}_w^{m,g} Q_h \frac{\rho}{\hat{\rho}}], \quad (31)$$

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

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

and similarly the variance  $\text{var}[\overline{Q}]$  is approximated as

$$\text{var}_h[Q_h] = \mathbb{E}[(\mathcal{S}_w^{m,g}[Q_h])^2] - \mathbb{E}[\mathcal{S}_w^{m,g}[Q_h]]^2 = \mathbb{E}_{\hat{\rho}}[(\mathcal{S}_w^{m,g}[Q_h])^2 \frac{\rho}{\hat{\rho}}] - \mathbb{E}_{\hat{\rho}}[\mathcal{S}_w^{m,g}[Q_h] \frac{\rho}{\hat{\rho}}]^2. \quad (32)$$

## 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, again under Assumption 5 that  $a \circ F$  and  $\hat{g}$  do not depend on  $\omega \in \Omega$ , and  $f$  is deterministic.

For notational simplicity we split the Jacobian as follows

$$\partial F(x, \omega) = I + \sum_{l=1}^{N_s} B_l(x) \sqrt{\mu_l} Y_l(\omega) + \sum_{l=N_s+1}^N B_l(x) \sqrt{\mu_l} Y_l(\omega). \quad (33)$$

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  $\Gamma_s$  (i.e.  $\mu_l = 0$  for  $l = N_s + 1, \dots, N$  in eqn (33)). A similar notation is used for the solution  $u(\mathbf{y}_s)$  as the restriction to  $\Gamma_s$  of  $u(\mathbf{y})$ , as well as 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 ( $|\text{var}[Q(\mathbf{y}_s, \mathbf{y}_f)] - \text{var}[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]|$ ) 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

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

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

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

or

$$\mathbb{E}[Q(\mathbf{y}_s, \mathbf{y}_f)^2 - Q(\mathbf{y}_s)^2] \leq \|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^1_\rho(\Gamma)} \|Q(\mathbf{y}_s, \mathbf{y}_f) + Q(\mathbf{y}_s)\|_{L^\infty(\Gamma)},$$

and

$$\begin{aligned} |\mathbb{E}[Q(\mathbf{y}_s, \mathbf{y}_f)]^2 - \mathbb{E}[Q(\mathbf{y}_s)]^2| &= |\mathbb{E}[Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)] \mathbb{E}[Q(\mathbf{y}_s, \mathbf{y}_f) + Q(\mathbf{y}_s)]| \\ &\leq \|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^1_\rho(\Gamma)} \|Q(\mathbf{y}_s, \mathbf{y}_f) + Q(\mathbf{y}_s)\|_{L^1_\rho(\Gamma)} \\ &\leq \|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)} \|Q(\mathbf{y}_s, \mathbf{y}_f) + Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)}. \end{aligned}$$

Therefore

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

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

$$\begin{aligned} & C_T \underbrace{\|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)}}_{\text{Truncation (I)}} + C_{FE} \underbrace{\|Q(\mathbf{y}_s) - Q_h(\mathbf{y}_s)\|_{L^1_\rho(\Gamma_s)}}_{\text{Finite Element (II)}} \\ & + C_{SG} \underbrace{\|Q_h(\mathbf{y}_s) - \mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]\|_{L^2_\rho(\Gamma_s)}}_{\text{Sparse Grid (III)}}, \end{aligned}$$

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  $Q : H_0^1(U) \rightarrow \mathbb{R}$  is a bounded linear functional then for any realization of  $\varphi(\mathbf{y}_s, \mathbf{y}_f)$  we have that

$$\begin{aligned} |Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)| &= |B(\mathbf{y}_s, \mathbf{y}_f; \varphi(\mathbf{y}_s, \mathbf{y}_f), \tilde{u}(\mathbf{y}_s, \mathbf{y}_f) - \tilde{u}(\mathbf{y}_s))| \\ &\leq a_{max} \mathbb{F}_{max}^d \mathbb{F}_{min}^{-2} \|\varphi(\mathbf{y}_s, \mathbf{y}_f)\|_{H_0^1(U)} \|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f) - \tilde{u}(\mathbf{y}_s)\|_{H_0^1(U)}, \end{aligned}$$

where we have denoted by  $\tilde{u}(\mathbf{y}_s) = (\tilde{u} \circ F)(\cdot, \mathbf{y}_s)$  and similarly  $\tilde{u}(\mathbf{y}_s, \mathbf{y}_f) = (\tilde{u} \circ F)(\cdot, \mathbf{y}_s, \mathbf{y}_f)$ . Following a similar argument as the proof from Lemma 2 we have that

$$\|\varphi(\mathbf{y}_s, \mathbf{y}_f)\|_{H_0^1(U)} \leq \frac{\|q\|_{L^2(\tilde{D})} C_P(U) \mathbb{F}_{max}^{d+2}}{a_{min} \mathbb{F}_{min}^d}$$

a.s., where  $q$  is defined in section 2.1.1. Thus

$$\|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)} \leq C_{TR} \|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f) - \tilde{u}(\mathbf{y}_s)\|_{L^2(\Gamma; H_0^1(U))},$$

where  $C_{TR} := a_{max} a_{min}^{-1} \mathbb{F}_{max}^{2d+2} \mathbb{F}_{min}^{-d-2} \|q\|_{L^2(\tilde{U})} C_P(U)$ . We now seek control on the error term  $e := \|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f) - \tilde{u}(\mathbf{y}_s)\|_{L^2_\rho(\Gamma; H_0^1(U))}$ . First we establish some notation and definitions. From Section 3 we have shown that the solution  $\tilde{u} \circ F$  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

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

then  $u \in C^0(\Gamma, H_0^1(U))$ . Furthermore, let

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

From Theorem 7 we have that  $\tilde{u} \circ F \in C^0(\Gamma; H_0^1(U)) \subset L^2_\rho(\Gamma; H_0^1(U))$ , thus  $\tilde{u} \circ F$  satisfies the following variational problem

$$\mathcal{A}(\tilde{u} \circ F, v) := E[B(\mathbf{y}_s, \mathbf{y}_f; \tilde{u} \circ F, v)] = E[\tilde{l}(\mathbf{y}_s, \mathbf{y}_f; v)] \quad \forall v \in L^2_\rho(\Gamma; H_0^1(U)).$$

The following lemma will be useful in deriving error estimates.

**Lemma 9.** For all  $w, v \in L^2_\rho(\Gamma; H_0^1(U))$  we have that

$$|\mathcal{A}(w, v)| \leq a_{\max} \mathbb{F}_{\max}^d \mathbb{F}_{\min}^{-2} \|w\|_{L^2_\rho(\Gamma; H_0^1(U))} \|v\|_{L^2_\rho(\Gamma; H_0^1(U))}.$$

PROOF.

$$\begin{aligned} |\mathcal{A}(w, v)| &\leq \sup_{\mathbf{y} \in \Gamma} \lambda_{\max}(G(\mathbf{y})) \mathbb{E} \left[ \int_U |\nabla w| |\nabla v| \right] \\ &\leq a_{\max} \mathbb{F}_{\max}^d \mathbb{F}_{\min}^{-2} \mathbb{E} [\|\nabla w\|_{L^2(U)} \|\nabla v\|_{L^2(U)}] \\ &\leq a_{\max} \mathbb{F}_{\max}^d \mathbb{F}_{\min}^{-2} \|w\|_{L^2_\rho(\Gamma; H_0^1(U))} \|v\|_{L^2_\rho(\Gamma; H_0^1(U))}. \end{aligned}$$

We can now derive the truncation error (I).

**Theorem 10.** Let  $\tilde{u}$  be the solution to the linear Problem 3 that satisfies Assumptions 1, 2, 3, 4 and 5. Furthermore, assume that  $\|(f \circ F)(\mathbf{y})\|_{W^{1,\infty}(U)}$  is bounded uniformly in  $\Gamma$  then

$$\|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f) - \tilde{u}(\mathbf{y}_s)\|_{L^2_\rho(\Gamma; H_0^1(U))} \leq \mathbb{C}_1 B_{\mathbb{T}} + \mathbb{C}_2 C_{\mathbb{T}},$$

where  $B_{\mathbb{T}} := \sup_{x \in U} \sum_{i=N_s+1}^N \sqrt{\mu_i} \|B_i(x)\|$ ,  $C_{\mathbb{T}} := \sum_{i=N_s+1}^N \sqrt{\mu_i} \|b_i(x)\|_{L^\infty(U)}$ ,

$$\begin{aligned} \mathbb{C}_1 &:= \mathcal{C} \left( C_P(U) \mathbb{F}_{\max}^{d-1} \mathbb{F}_{\min}^{-2} d \|f \circ F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(\Gamma; L^2(U))} \right. \\ &\quad \left. + a_{\max} H(\mathbb{F}_{\max}, \mathbb{F}_{\min}, \tilde{\delta}, d) \|\hat{\mathbf{w}}\|_{L^2(\Gamma; H_0^1(U))} \right. \\ &\quad \left. + a_{\max} H(\mathbb{F}_{\max}, \mathbb{F}_{\min}, \tilde{\delta}, d) \|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2_\rho(\Gamma; H_0^1(U))} \right), \\ \mathbb{C}_2 &:= \mathcal{C} \|\hat{v}\|_{L^\infty(U)} \mathbb{F}_{\max}^d C_P(U) \|f\|_{L^\infty(\Gamma; W^{1,\infty}(U))} \|\chi_U\|_{L^2(U)}, \end{aligned}$$

$\mathcal{G} := \cup_{\omega \in \Omega} D(\omega)$ ,  $\mathcal{C} := \frac{C_P(U)^2}{a_{\min} \mathbb{F}_{\min}^d \mathbb{F}_{\max}^{-2}}$ , and  $H(\mathbb{F}_{\max}, \mathbb{F}_{\min}, \tilde{\delta}, d) := \mathbb{F}_{\max}^{d-1} \mathbb{F}_{\min}^{-3} (\mathbb{F}_{\max} (2 + \mathbb{F}_{\min}^{-1} (1 - \tilde{\delta})) + \mathbb{F}_{\min}^{-1} d)$ .

PROOF. We follow a similar strategy as in [8, 19] to compute the bounds for the truncation of the stochastic variables to  $\Gamma_s$ . Consider the solution to Problem 3  $\tilde{u}(\mathbf{y}_s) \in C^0(\Gamma_s; H_0^1(U)) \subset L^2_\rho(\Gamma; H_0^1(U))$ , where the matrix of coefficients  $G(\mathbf{y}_s)$  depends only on the variables  $Y_1, \dots, Y_{N_s}$ ,

$$\mathbb{E}[B(\mathbf{y}_s; \tilde{u}(\mathbf{y}_s), v)] = \mathbb{E}[\tilde{l}(\mathbf{y}_s; v)] \quad \forall v \in L^2_\rho(\Gamma_s; H_0^1(U)).$$

Furthermore the variational form is still valid  $\forall v \in L^2_\rho(\Gamma; H_0^1(U))$  i.e.

$$\mathcal{A}_{N_s}(\tilde{u}(\mathbf{y}_s), v) := \mathbb{E}[B(\mathbf{y}_s; \tilde{u}(\mathbf{y}_s), v)] = \mathbb{E}[\tilde{l}(\mathbf{y}_s; v)] \quad \forall v \in L^2_\rho(\Gamma; H_0^1(U)).$$

Now, Observe that  $\forall v \in L^2_\rho(\Gamma; H_0^1(U))$  we have that

$$\begin{aligned} \mathcal{A}_{N_s}(v, v) &\geq \inf_{\mathbf{y}_s \in \Gamma_s} \lambda_{\min}(G(\mathbf{y}_s)) \mathbb{E}[\|\nabla v\|_{L^2(U)}^2] \\ &\geq a_{\min} \mathbb{F}_{\min}^d \mathbb{F}_{\max}^{-2} \mathbb{E}[\|\nabla v\|_{L^2(U)}^2] \\ &\geq a_{\min} \mathbb{F}_{\min}^d \mathbb{F}_{\max}^{-2} C_P(U)^{-2} \|v\|_{L^2_\rho(\Gamma; H_0^1(U))}^2. \end{aligned}$$



By adapting the proof from Strang's Lemma and applying Lemma 9 we have that for all  $v \in L^2_\rho(\Gamma; H^1_0(U))$

$$\begin{aligned} \|\tilde{u}(\mathbf{y}_s) - v\|_{L^2_\rho(\Gamma; H^1_0(U))}^2 &\leq \mathcal{C}(\mathcal{A}_{N_s}(\tilde{u}(\mathbf{y}_s) - v, \tilde{u}(\mathbf{y}_s) - v) \pm \mathcal{A}(\tilde{u}(\mathbf{y}) - v, \tilde{u}(\mathbf{y}_s) - v)) \\ &\leq \mathcal{C}(a_{max}\mathbb{F}_{max}^d\mathbb{F}_{min}^{-2}\|\tilde{u}(\mathbf{y}) - v\|_{L^2_\rho(\Gamma; H^1_0(U))}\|\tilde{u}(\mathbf{y}_s) - v\|_{L^2_\rho(\Gamma; H^1_0(U))}) \\ &\quad + |\mathcal{A}(v, \tilde{u}(\mathbf{y}_s) - v) - \mathcal{A}_{N_s}(v, \tilde{u}(\mathbf{y}_s) - v)| \\ &\quad + |\mathbb{E}[\tilde{l}(\mathbf{y}_s; \tilde{u}(\mathbf{y}_s) - v) - \tilde{l}(\mathbf{y}_s, \mathbf{y}_f; \tilde{u}(\mathbf{y}_s) - v)]|. \end{aligned}$$

Now, pick  $v = \tilde{u}(\mathbf{y}_s, \mathbf{y}_f)$ , thus

$$\|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f) - \tilde{u}(\mathbf{y}_s)\|_{L^2_\rho(\Gamma; H^1_0(U))} \leq \mathcal{C}(\mathcal{B}_1 + \mathcal{B}_2),$$

where

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

and

$$\mathcal{B}_2 := \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} |\mathbb{E}[\tilde{l}(\mathbf{y}_s; \tilde{u}(\mathbf{y}_s) - w) - \tilde{l}(\mathbf{y}_s, \mathbf{y}_f; \tilde{u}(\mathbf{y}_s) - w)]|.$$

**Bound for  $\mathcal{B}_1$ :** For notational simplicity we rewrite (8) as

$$\partial F(\mathbf{y}_s, \mathbf{y}_f) = I + A_{N_s}^s(\mathbf{y}_s) + A_{N_f}^f(\mathbf{y}_f)$$

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

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

$$\|G(\mathbf{y}_s, \mathbf{y}_f) - G(\mathbf{y}_s)\| \leq a_{max} \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\|. \quad (34)$$

Now,  $\forall x \in U$  and  $\forall \mathbf{y} \in \Gamma$

$$\begin{aligned} \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\| &= \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s) \pm |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \\ &\leq |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \| |\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T} \| \\ &\quad + \| |\partial F(\mathbf{y}_s, \mathbf{y}_f)| - |\partial F(\mathbf{y}_s)| \| \| \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T} \| \\ &\leq \mathbb{F}_{max}^d \| \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T} \| \\ &\quad + \mathbb{F}_{min}^{-2} \| |\partial F(\mathbf{y}_s, \mathbf{y}_f)| - |\partial F(\mathbf{y}_s)| \|. \end{aligned} \quad (35)$$

Applying the matrix identity  $(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}$  where  $A = I + A_{N_s}^s(\mathbf{y}_s)$ ,  $B = -A_{N_f}^f(\mathbf{y}_f)$  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

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

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,  $\forall x \in U$  and  $\forall \mathbf{y} \in \Gamma$

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

It follows that  $\forall x \in U$  and  $\forall \mathbf{y} \in \Gamma$

$$\|\partial F(\mathbf{y}_s, \mathbf{y}_f)^{-1} \partial F(\mathbf{y}_s, \mathbf{y}_f)^{-T} - \partial F(\mathbf{y}_s)^{-1} \partial F(\mathbf{y}_s)^{-T}\| \leq B_{\mathbb{T}} \mathbb{F}_{min}^{-3} (2 + \mathbb{F}_{min}^{-1} (1 - \tilde{\delta})). \quad (36)$$

From Theorem 2.12 in [20] ( $A, E \in \mathbb{C}^{d \times d}$  then  $|\det(A + E) - \det(A)| \leq d \|E\| \max\{\|A\|, \|A + E\|\}^{d-1}$ ) we obtain  $\forall x \in U$  and  $\forall \mathbf{y} \in \Gamma$

$$\left| |\partial F(\mathbf{y}_s, \mathbf{y}_f)| - |\partial F(\mathbf{y}_s)| \right| \leq \mathbb{F}_{max}^{d-1} \mathbb{F}_{min}^{-2} B_{\mathbb{T}} d. \quad (37)$$

Combining equations (35), (36) and (37) we obtain

$$\sup_{x \in U, \mathbf{y} \in \Gamma} \|J(\mathbf{y}_s, \mathbf{y}_f) - J(\mathbf{y}_s)\| \leq B_{\mathbb{T}} H(\mathbb{F}_{max}, \mathbb{F}_{min}, \tilde{\delta}, d).$$

Thus,

$$\mathcal{B}_1 \leq a_{max} B_{\mathbb{T}} H(\mathbb{F}_{max}, \mathbb{F}_{min}, \tilde{\delta}, d) \|\tilde{u}(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2_\rho(\Gamma; H^1_\rho(U))}.$$

**Bound for  $\mathcal{B}_2$ :**

$$\begin{aligned} \mathcal{B}_2 &\leq \sup_{w \in L^2_\rho(\Gamma; H^1_\rho(U))} \|w\|_{L^2_\rho(\Gamma; H^1_\rho(U))}^{-1} \left| \mathbb{E} \left[ \int_U (f \circ F(\mathbf{y}_s)) |\partial F(\mathbf{y}_s)| \right. \right. \\ &\quad \left. \left. - f \circ F(\mathbf{y}_s, \mathbf{y}_f) |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \right] w \right. \\ &\quad \left. + \int_U (\nabla \hat{\mathbf{w}})^T (G(\mathbf{y}_s) - G(\mathbf{y}_s, \mathbf{y}_f)) \nabla w \right| \\ &\leq \sup_{w \in L^2_\rho(\Gamma; H^1_\rho(U))} \|w\|_{L^2_\rho(\Gamma; H^1_\rho(U))}^{-1} \mathbb{E} \left[ \int_U |(f \circ F(\mathbf{y}_s)) - (f \circ F(\mathbf{y}_s, \mathbf{y}_f))| |\partial F(\mathbf{y}_s)| |w| \right. \\ &\quad \left. + |f \circ F(\mathbf{y}_s, \mathbf{y}_f)| (|\partial F(\mathbf{y}_s)| - |\partial F(\mathbf{y}_s, \mathbf{y}_f)|) |w| \right. \\ &\quad \left. + |(\nabla \hat{\mathbf{w}})^T (G(\mathbf{y}_s) - G(\mathbf{y}_s, \mathbf{y}_f)) \nabla w| \right]. \end{aligned} \quad (38)$$

Now we have that

$$\begin{aligned}
& \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} \mathbb{E} \left[ \int_U |(f \circ F(\mathbf{y}_s) - f \circ F(\mathbf{y}_s, \mathbf{y}_f))| |\partial F(\mathbf{y}_s)| |w| \right] \\
& \leq \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} \mathbb{E} \left[ \max_{x \in U} |\partial F(\mathbf{y}_s)| \|f \circ F(\mathbf{y}_s) - f \circ F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(U)} \|w\|_{L^2(U)} \right] \\
& \leq \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} \mathbb{F}_{max}^d \mathbb{E} \left[ \|f \circ F(\mathbf{y}_s) - f \circ F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(U)} \right. \\
& \quad \left. C_P(U) \|w\|_{H^1_0(U)} \right] \\
& \leq \mathbb{F}_{max}^d C_P(U) \mathbb{E} \left[ \|f(\cdot, \mathbf{y})\|_{W^{1,\infty}(\mathcal{G})} \|F(\mathbf{y}_s) - F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(U)} \right] \\
& \leq \mathbb{F}_{max}^d C_P(U) \|f\|_{L^\infty(\Gamma; W^{1,\infty}(\mathcal{G}))} \|\chi_U\|_{L^2(U)} \sup_{\mathbf{y} \in \Gamma, x \in U} |F(\mathbf{y}_s) - F(\mathbf{y}_s, \mathbf{y}_f)|,
\end{aligned} \tag{39}$$

where  $\chi_U$  is the characteristic function defined on  $U$ . Note that  $\forall x \in U$  and  $\forall \mathbf{y} \in \Gamma$

$$\begin{aligned}
& \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} \mathbb{E} \left[ \int_U |f \circ F(\mathbf{y}_s, \mathbf{y}_f)| (|\partial F(\mathbf{y}_s)| - |\partial F(\mathbf{y}_s, \mathbf{y}_f)|) |w| \right] \\
& \leq \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} \mathbb{E} \left[ \int_U \|f \circ F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(U)} \|w\|_{L^2(U)} \right] \\
& \quad \sup_{\mathbf{y} \in \Gamma, x \in U} \left| |\partial F(\mathbf{y}_s)| - |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \right| \\
& \leq C_P(U) \|f \circ F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(\Gamma; L^2(U))} \sup_{\mathbf{y} \in \Gamma, x \in U} \left| |\partial F(\mathbf{y}_s)| - |\partial F(\mathbf{y}_s, \mathbf{y}_f)| \right|,
\end{aligned} \tag{40}$$

and

$$\begin{aligned}
& \sup_{w \in L^2_\rho(\Gamma; H^1_0(U))} \|w\|_{L^2_\rho(\Gamma; H^1_0(U))}^{-1} \mathbb{E} \left[ (\nabla \hat{\mathbf{w}})^T (G(\mathbf{y}_s) - G(\mathbf{y}_s, \mathbf{y}_f)) \nabla w \right] \\
& \leq \|\hat{\mathbf{w}}\|_{L^2(\Gamma; H^1_0(U))} \sup_{\mathbf{y} \in \Gamma, x \in U} \|G(\mathbf{y}) - G(\mathbf{y}_s)\|.
\end{aligned} \tag{41}$$

Substituting (34), (37), (39), (40), and (41) in (38) we obtain

$$\begin{aligned}
\mathcal{B}_2 & \leq B_{\mathbb{T}} \left( C_P(U) \mathbb{F}_{max}^{d-1} \mathbb{F}_{min}^{-2} d \|f \circ F(\mathbf{y}_s, \mathbf{y}_f)\|_{L^2(\Gamma; L^2(U))} \right. \\
& \quad \left. + a_{max} H(\mathbb{F}_{max}, \mathbb{F}_{min}, \tilde{\delta}, d) \|\hat{\mathbf{w}}\|_{L^2(\Gamma; H^1_0(U))} \right) \\
& \quad + C_{\mathbb{T}} \|\hat{v}\|_{L^\infty(U)} \mathbb{F}_{max}^d C_P(U) \|f\|_{L^\infty(\Gamma; W^{1,\infty}(\mathcal{G}))} \|\chi_U\|_{L^2(U)}.
\end{aligned}$$

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  $H_h(U) \subset H^1_0(U)$  and the regularity of the solution  $u$ . From (25) we obtain the following bound:

$$\|\tilde{u}(\mathbf{y}_s) - u_h(\mathbf{y}_s)\|_{L^2_\rho(\Gamma_s; H^1_0(U))} \leq C_{\Gamma_s}(r) h^r$$

for some constant  $r \in \mathbb{N}$  and  $C_{\Gamma_s}(r) := \int_{\Gamma_s} C(r, \tilde{u}(\mathbf{y}_s)) \rho(\mathbf{y}_s) d\mathbf{y}_s$ . 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

$$\|\varphi(\mathbf{y}_s) - \varphi_h(\mathbf{y}_s)\|_{L^2_{\hat{\rho}}(\Gamma_s; H^1_0(U))} \leq D_{\Gamma_s}(r) h^r,$$

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

$$\mathbb{E} [|Q(\tilde{u}(\mathbf{y}_s) - u_h(\mathbf{y}_s))|] \leq a_{max} \mathbb{F}_{max}^d \mathbb{F}_{min}^{-2} C_{\Gamma_s}(r) D_{\Gamma_s}(r) h^{2r}. \quad (42)$$

### 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

$$\|Q(u_h(\mathbf{y}_s)) - Q(\mathcal{S}_w^{m,g}[u_h(\mathbf{y}_s)])\|_{L^2_{\hat{\rho}}(\Gamma_s)} \leq a_{max} \mathbb{F}_{max}^d \mathbb{F}_{min}^{-2} \|e\|_{L^2_{\hat{\rho}}(\Gamma_s; H^1_0(U))},$$

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

$$\|e\|_{L^2_{\hat{\rho}}(\Gamma_s; H^1_0(U))} \leq \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma_s)} \|e\|_{L^2_{\rho}(\Gamma_s; H^1_0(U))}.$$

The error term  $\|e\|_{L^2_{\hat{\rho}}(\Gamma_s; H^1_0(U))}$  is controlled by the number of collocation knots  $\eta$  (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 7 the solution  $\tilde{u}(\mathbf{y}_s)$  admits an extension in  $\mathbb{C}^{N_s}$  i.e.  $\mathbf{y}_s \rightarrow \mathbf{z}_s \in \mathbb{C}^{N_s}$  and  $\tilde{u}(\mathbf{z}_s) \in C^0(\Theta_\beta; H^1_0(U))$ . 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  $\Theta_\beta$  and  $u_h(\mathbf{z}_s) \in C^0(\Theta_\beta; H_h(U))$ .

In [17, 21] the authors derive error estimates for isotropic and anisotropic Smolyak sparse grids with Clenshaw-Curtis and Gaussian abscissas where  $\|e\|_{L^2_{\hat{\rho}}(\Gamma_s; H^1_0(U))}$  exhibit algebraic or subexponential convergence with respect to the number of collocation knots  $\eta$  (See Theorems 3.10, 3.11, 3.18 and 3.19 in [21] for more details). However, for these estimates to be valid the solution  $u$  has to admit an extension on a polyellipse in  $\mathbb{C}^{N_s}$ ,  $\mathcal{E}_{\sigma_1, \dots, \sigma_{N_s}} := \prod_{i=1}^{N_s} \mathcal{E}_{n, \sigma_n}$ , where

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

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

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

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

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

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

From Theorem 3.11 [21], given a sufficiently large  $\eta$  for a nested CC sparse grid we obtain the following estimate

$$\|e\|_{L_\beta^2(\Gamma_s; H_0^1(U))} \leq \mathcal{Q}(\sigma, \delta^*, N_s) \eta^{\mu_3(\sigma, \delta^*, N_s)} \exp\left(-\frac{N_s \sigma}{2^{1/N_s}} \eta^{\mu_2(N_s)}\right), \quad (43)$$

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  $\delta^*$  are defined in [21] equations (3.11) and (3.12).

## 6. Complexity and Tolerance

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

Let  $N_h$  be the number of degrees of freedom to solve the semi-discrete approximation  $u_h \in H_h(U) \subset H_0^1(U)$ . We assume that the complexity for solving for  $u_h$  is  $\mathcal{O}(N_h^q)$ , where the constant  $q \geq 1$  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  $\mathcal{S}_w^{m,g}$  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  $\mathbb{E}[(\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)])^2] - \mathbb{E}[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]^2$  and the mean term  $\mathbb{E}[\mathcal{S}_w^{m,g}[Q_h(\mathbf{y}_s)]]$  with respect to a given user tolerance is

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

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  $\|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L_\beta^2(\Gamma)} \leq \frac{\text{tol}}{3C_T}$  with respect to the decay of  $\mu_i$ . First, make the assumption

that  $B_{\mathbb{T}} = \sup_{x \in U} \sum_{i=N_s+1}^N \sqrt{\mu_i} \|B_i(x)\|_2 \leq C_D N_s^{-l}$  for some uniformly bounded  $C_D > 0$ . Furthermore, assume that  $\|b_i(x)\|_{L^\infty(U)} \leq D_D \sup_{x \in U} \|B_i(x)\|_2$  for  $i = 1, \dots, N$  where  $D_D > 0$  is uniformly bounded, thus  $C_{\mathbb{T}} = \sup_{x \in U} \sum_{i=N_s+1}^N \sqrt{\mu_i} \|b_i(x)\|_2 \leq C_D D_D N_s^{-l}$ . It follows that  $\|Q(\mathbf{y}_s, \mathbf{y}_f) - Q(\mathbf{y}_s)\|_{L^2_\rho(\Gamma)} \leq \frac{tol}{3C_{\mathbb{T}}}$  if

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

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

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

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

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

then  $\|Q(\mathbf{y}_s) - Q_h(\mathbf{y}_s)\|_{L^2_\rho(\Gamma; H_0^1(U))} \leq \frac{tol}{3C_{FE}}$ . Assuming that  $N_h$  grows as  $\mathcal{O}(h^{-d})$  then

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

for some constant  $D_3 > 0$ .

(c) **Sparse Grid:** Following the same strategy as in [21] (equation (3.39)), to simplify the bound (43) choose  $\delta^* = (e \log(2) - 1) / \tilde{C}_2(\sigma)$  and  $\tilde{C}_2(\sigma)$ . Thus  $\|e\|_{L^2_\rho(\Gamma_s; H_0^1(U))} \leq$

$$\frac{tol}{3C_{SG} C_{\mathbb{T}}} \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma_s)}^{-1} \text{ if}$$

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

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

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

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:

$$\begin{aligned} F(x_1, x_2) &= (x_1, x_2) + e(x_1, \omega)(0, x_2 - 0.5) & \text{if } x_2 > 0.5 \\ F(x_1, x_2) &= (x_1, x_2) & \text{if } 0 \leq x_2 \leq 0.5 \end{aligned}$$

for some positive constant  $c > 0$ . In other words we deform only the upper half of the domain and fix the bottom half. We set the Dirichlet boundary conditions to zero everywhere except at the upper border to  $\tilde{u}(x_1, x_2)|_{\partial U} = g(x_1)$ , where  $g(x_1) := \exp\left(\frac{-1}{1-4(x_1-0.5)^2}\right)$  (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

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

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

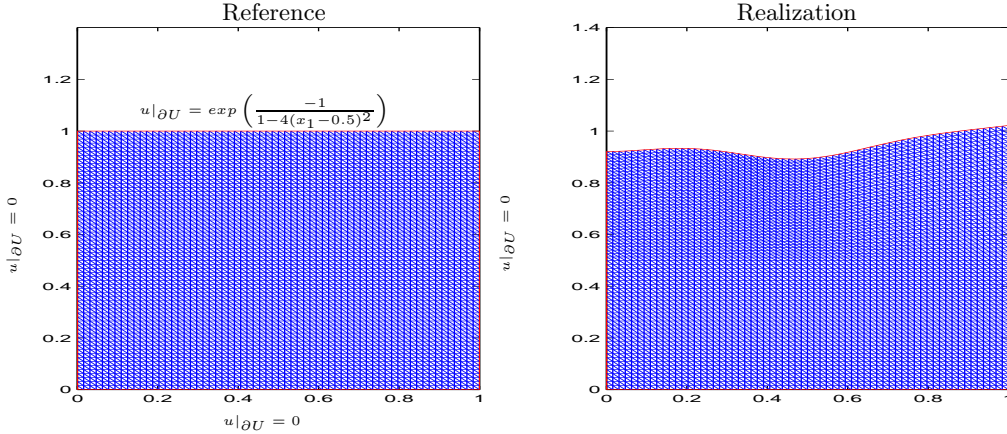


Figure 2: Stochastic deformation of a square domain. (left) Reference square domain with Dirichlet boundary conditions. (right) Vertical deformation from stochastic model.

**Numerical Experiment 1:** For this numerical experiment we compute the QoI with a multi-linear Clenshaw-Curtis sparse grid [22, 23]. The parameters of the experiments are as follows:

- (a)  $a(x) = 1$  for all  $x \in U$ .
- (b) Stochastic Model. We split  $e(x_1, \omega)$  into large and small perturbations as  $e(x_1, \omega) := e_S(x_1, \omega) + e_F(x_1, \omega)$ , where

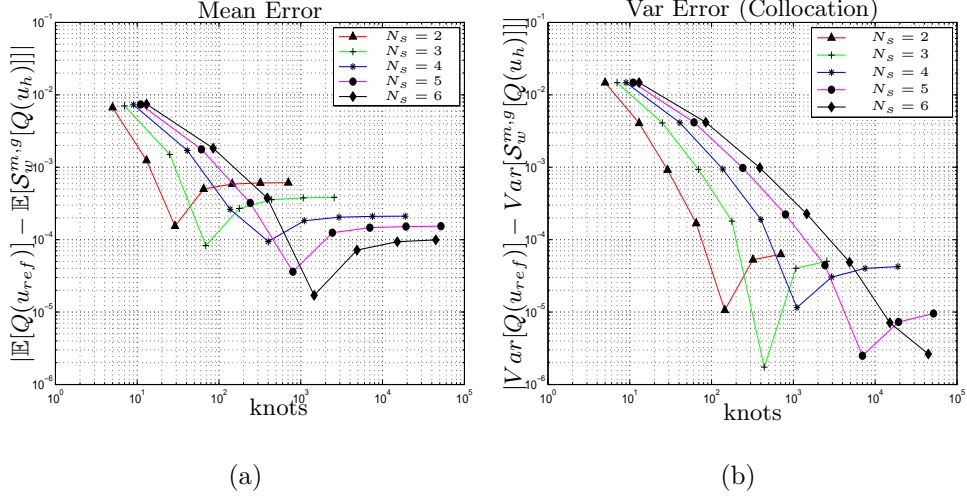


Figure 3: Collocation results (Clenshaw-Curtis multi-linear sparse grid) for  $N_s = 2, \dots, 6$  with linear decay. (a) Mean error with respect to the reference solution. Observe that the convergence rate decays subexponentially until truncation saturation is reached. (b) Variance error with respect to reference solution. For this case we also observe that the convergence rate is faster than polynomial.

$$e_S(x_1, \omega) := cY_1(\omega) \left( \frac{\sqrt{\pi}L}{2} \right)^{1/2} + c \sum_{n=2}^{N_s} \sqrt{\mu_n} \varphi_n(x_1) Y_n(\omega);$$

$$e_F(x_1, \omega) := c \sum_{n=N_s+1}^N \sqrt{\mu_n} \varphi_n(x_1) Y_n(\omega).$$

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

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

Thus for  $n > 1$  we have that

$$B_n = \begin{bmatrix} 0 & 0 \\ c(x_2 - 0.5) \partial_{x_1} \varphi_n(x_1) & 0 \end{bmatrix}.$$

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)  $L = 1/2$ ,  $L_P = 1$ ,  $c = 0.1533$ ,  $N = 15$ .

(f)  $257 \times 257$  triangular mesh.

(g)  $\mathbb{E}[Q_h]$  and  $\text{var}[Q_h]$ , are computed with a Clenshaw-Curtis isotropic sparse grid (Sparse Grid Toolbox V5.1, [22, 23]).



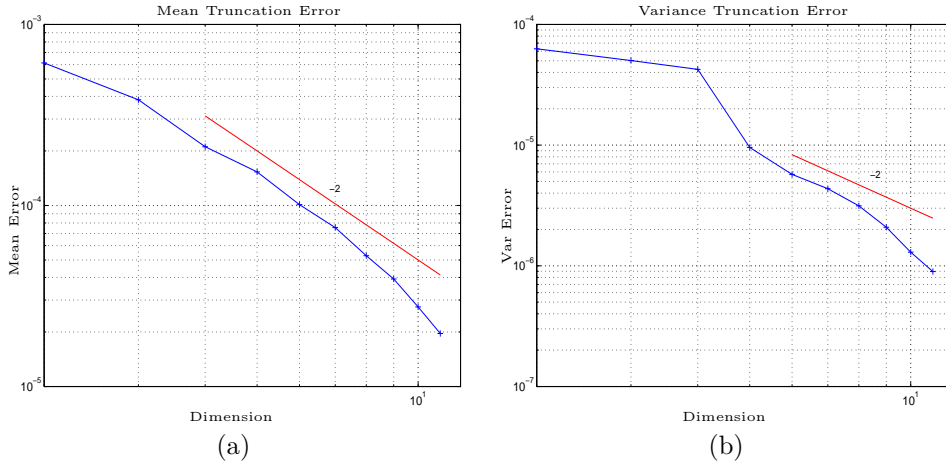


Figure 4: Truncation Error with respect to the number of dimensions. (a) Mean error. (b) Variance error. In both cases the decay appears faster than linear, which is faster than the predicted convergence rate.

- (h) The reference solutions  $\text{var}[Q_h(u_{ref})]$  and  $\mathbb{E}[Q_h(u_{ref})]$  are computed with a multi-linear adaptive Sparse Grid ( $\approx 30,000$  knots) [24] with a  $257 \times 257$  mesh for  $N = 15$  dimensions.
- (i) The QoI is normalized by the reference solution  $Q(U)$ .

In Figure 3 we show the results of the matlab code for the truncated dimensions  $N_s = 2, \dots, 6$  and compare the results with respect to a  $N = 15$  dimensional adaptive sparse grid method collocation with  $\approx 30,000$  collocation points [24]. 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  $N_s = 2, \dots, 6$ . For (a) notice the subexponential 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 despite the use of a multi-linear sparse grid, whose performance is less than higher order Lagrange polynomial sparse grid shown in the second numerical experiment.

We now analyze the decay of the truncation error. For  $N_s = 2, \dots, 5$  we compute the mean and variance error as in (g). However, for  $N_s = 6, \dots, 11$  a dimension adaptive sparse grid with 15,000 to 30,000 sparse grid points is used instead to compute the mean and variance. This should be enough to make the error contribution from the 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.

**Numerical Experiment 2:** For this numerical experiment we compute the QoI

with a higher order Lagrange polynomial sparse grid with Chebyshev Gauss-Lobatto abscissas [22, 23]. For this experiment we show the convergence rates for the variance and sparse grid error only. The setup for the numerical experiment is changed as follows:

- (a) Since the convergence is much faster than the multi-linear sparse grid we increase the mesh size to a  $1024 \times 1024$  triangular mesh.
- (b)  $\text{var}[Q_h]$  is computed for  $N_s = 2, \dots, 5$  with the Chebyshev Gauss-Lobatto abscissas.
- (c) As a comparison  $\text{var}[Q_h]$  is also computed for  $N_s = 2, \dots, 5$  with the dimension adaptive multi-linear sparse grid with 15,000 to 20,000 adaptive nodes.

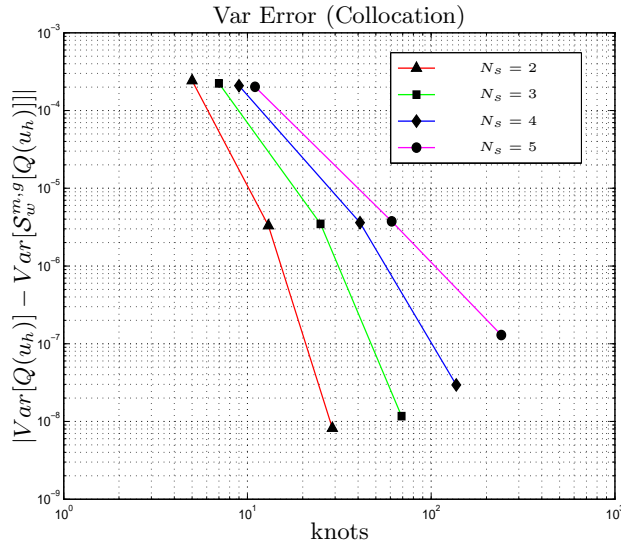


Figure 5: Collocation results for higher order Lagrange polynomial sparse grid with Chebyshev Gauss-Lobatto abscissas. Due to the higher order Lagrange interpolants the convergence is significantly faster compared to the multi-linear sparse grid in Figure 3

In Figure 5 the results for  $|\text{Var}[Q(u)] - \text{Var}[\mathcal{S}_w^{m,g}[Q(u)]]|$  for  $N_s = 2, \dots, 5$  are plotted. Due to the higher order Lagrange interpolants the convergence is significantly faster compared to the multi-linear sparse grid in Figure 3. Note that we did not do a direct comparison with  $|\text{Var}[Q(u_{ref})] - \text{Var}[\mathcal{S}_w^{m,g}[Q(u_h)]]|$  as the convergence rate was so fast that the Truncation error was reached by the second level of the sparse grid ( $w = 2$ ).

## 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” 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  $\Theta_\beta$  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  $\approx 0.17$ ). This is a clear advantage over the perturbation approaches that are restricted to small deviations. In addition, the numerical experiments confirm the subexponential 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 [17].

## Acknowledgements

We appreciate the advice from Quan Long and Serge Prudhomme. We are also appreciative of the efforts by the reviewers. In particular, one reviewer provided a wealth of corrections, comments and advice.

## 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 electro-magnetics, *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] R. A. Adams, *Sobolev Spaces*, Academic Press, 1975.
- [7] L. C. Evans, *Partial Differential Equations*, Vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, Rhode Island, 1998.
- [8] 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.
- [9] F. Jabbari, Chapter 3: Eigenvalues, singular values and pseudo inverse, University of California Irvine, Irvine, California.  
URL <http://gram.eng.uci.edu/~fjabbari/me270b/me270b.html>
- [10] I. C. F. Ipsen, D. J. Lee, Determinant approximations, ArXiv (arXiv:1105.0437v1).
- [11] D. London, A note on matrices with positive definite real part, *Proceedings of the American Mathematical Society* 82 (3) (1981) pp. 322–324.
- [12] S. G. Krantz, *Function Theory of Several Complex Variables*, AMS Chelsea Publishing, Providence, Rhode Island, 1992.
- [13] R. Gunning, H. Rossi, *Analytic Functions of Several Complex Variables*, American Mathematical Society, 1965.

- [14] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, *Soviet Mathematics, Doklady* 4 (1963) 240–243.
- [15] V. Barthelmann, E. Novak, K. Ritter, High dimensional polynomial interpolation on sparse grids, *Advances in Computational Mathematics* 12 (2000) 273–288.
- [16] 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.
- [17] 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.  
URL <http://epubs.siam.org/doi/pdf/10.1137/070680540>
- [18] A. Chkifa, A. Cohen, C. Schwab, High-dimensional adaptive sparse polynomial interpolation and applications to parametric pdes, *Foundations of Computational Mathematics* 14 (4) (2014) 601–633. doi:10.1007/s10208-013-9154-z.  
URL <http://dx.doi.org/10.1007/s10208-013-9154-z>
- [19] 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.  
URL <http://epubs.siam.org/doi/pdf/10.1137/S0036142902418680>
- [20] I. Ipsen, R. Rehman, Perturbation bounds for determinants and characteristic polynomials, *SIAM Journal on Matrix Analysis and Applications* 30 (2) (2008) 762–776.  
URL <http://epubs.siam.org/doi/pdf/10.1137/070704770>
- [21] 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.  
URL <http://epubs.siam.org/doi/pdf/10.1137/060663660>
- [22] A. Klimke, B. Wohlmuth, Algorithm 847: spinterp: Piecewise multilinear hierarchical sparse grid interpolation in MATLAB, *ACM Transactions on Mathematical Software* 31 (4).
- [23] A. Klimke, *Sparse Grid Interpolation Toolbox – user’s guide*, Tech. Rep. IANS report 2007/017, University of Stuttgart (2007).
- [24] T. Gerstner, M. Griebel, Dimension-adaptive tensor-product quadrature, *Computing* 71 (1) (2003) 65–87.