Masterarbeit

A review of regularity, sparse approximation and quadrature results for stochastic partial differential equations

Sophie Madeleine Knell

October 25, 2015

1. Betreuer: Prof. Dr. Volker John

2. Betreuer: Dr. Martin Eigel

Fachbereich Mathematik und Informatik

Freie Universität Berlin

Contents

1 Introduction

Physical and engineering problems are often modelled by partial differential equations (PDEs) and solved by numerical simulations. However, uncertainty in the available data in these problems such as coefficients, forcing terms, boundary conditions and geometry affects the results of the computer simulations and contributes to a discrepancy in the simulation outputs and the observations (see [14, p. 524]). In the last few years research about how uncertainty in the available data affects the outputs has grown. There are a variety of methods for solving the arising stochastic PDEs, for example spectral Galerkin approximations or the stochastic collocation method. Here, the stochastic collocation method will be considered as the numerical procedure for solving a stochastic PDE. As in spectral Galerkin approximations, the spatial variables are approximated by standard approximations such as finite element methods, and the stochastic variable is approximated by polynomials. Its advantage, in contrast to the spectral Galerkin approximation, is a decoupling of spatial and stochastic variables. The stochastic PDE in consideration here is an elliptic problem describing diffusion in a stationary system.

The organisation of the thesis is the following. In Section 2 the elliptic problem and its weak formulation are given, and its parametric formulation is derived.

In Section 3 the focus is on the diffusion problem and its mixed form. The mixed form is worthwhile considering because it is of interest in applications. For instance, in modelling a groundwater flow problem the flux is of importance and by the mixed form a more accurate approximation may be achieved. The diffusion problem and its mixed form are subsequently defined for a random force term and diffusion coefficient. Existence, uniqueness and the regularity of a weak solution are investigated in the case of a uniformly bounded diffusion coefficient as well as of a diffusion coefficient bounded by random variables only. The latter poses several significant difficulties in theory and is highly relevant for applications, in particular with lognormal fields. This thesis aims to give an overview of and to consistently present existing results on the diffusion problem and its mixed form. Some remaining gaps are filled as to give an entire review for the diffusion problem and its mixed form with uniformly diffusion coefficients and diffusion coefficients bounded by random variables. In the literature, for example in [1], [7] and [8], results about existence and uniqueness of a weak solution of the diffusion problem can be found for the case of a uniformly bounded diffusion coefficient. In addition, results about the analytic extension to the complex domain are given. In this thesis the results are derived by the same procedure as in [11], where only the diffusion coefficient bounded by random variables and the mixed form is examined. Two different analyticity results are given. These are subsequently used to derive error estimates of the stochastic collocation method with a tensor product or a sparse grid interpolation. To derive additional regularity statements of the diffusion problem with uniformly bounded diffusion coefficient [2] is consulted. They are required in the multilevel approach in Section 5. The results for the other relevant problem settings are then derived in a similar manner.

In Section 4, the collocation method is described by approximations on tensor product as well as on sparse grids and it is related to numerical quadrature. Existing theorems on error estimates for the tensor product grid and the Smolyak sparse grid are given and discussed. These can be found in the literature in [1], [15] and [22].

The multilevel approximation and quadrature is introduced in Section 5 as another kind of sparse grid method, for which the results of the previous sections can be reused. The error estimates given for the multilevel approximation of the diffusion problem's weak solution are according to [24]. They are stated for the quadrature in the same manner. Likewise, results for the mixed form are derived.

Section 6 gives an outlook and conclusions.

In the Appendices B and C basics of functional analysis and of analytic functions needed in this thesis can be found. Appendix D contains a list of notations, mainly those appearing in different sections of the thesis.

2 The elliptic problem and its parametric formulation

The problem setting is given in this section. In the first part the problem is described on an infinite-dimensional probability space. Then, it is reduced to an N-dimensional probability space by introducing N independent random variables as “coordinates”. The parametrization is an important step in approximating the solution numerically.
A formulation is given which is as general as possible in order to cover a wider range of problems. Such problems are considered in Section 3.

2.1 Elliptic problem

Let $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a spatial domain with Lipschitz-boundary ∂D. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space (i.e., for all $A \subset B$ with $B \in \mathcal{F}$ and $\mathbb{P}(B) = 0$ it follows $A \in \mathcal{F}$) with

- Ω being the set of events or outcomes $\omega \in \Omega$,
- $\mathcal{F} \subset 2^\Omega$ being the σ-algebra of outcomes and
- $\mathbb{P} : \mathcal{F} \to [0, 1]$ being a probability measure.

Let

$$\begin{aligned} a : \Omega \times D &\to \mathbb{R} \\ (\omega, \mathbf{x}) &\mapsto a(\omega, \mathbf{x}) \end{aligned}$$

be a random coefficient with $\omega \in \Omega$ and $\mathbf{x} \in D$. Similarly, a force term $\mathbf{f}$ is defined by

$$\begin{aligned} \mathbf{f} : \Omega \times D &\to \mathbb{R}^m \\ (\omega, \mathbf{x}) &\mapsto \mathbf{f}(\omega, \mathbf{x}), \end{aligned}$$

where $m \in \mathbb{N}$. A solution $\mathbf{w} : \Omega \times \bar{D} \to \mathbb{R}^n$ is sought with $n \in \mathbb{N}$ such that $\mathbb{P}$-almost everywhere ($\mathbb{P}$-a.e.) in Ω (i.e., the set of elements in Ω for which the equations do not hold has measure zero) it holds

$$\mathcal{L}(a)(\mathbf{w}) = \mathbf{f} \text{ in } D \tag{2.1}$$

with additional equations for suitable boundary conditions, where $\mathcal{L}$ is a (possibly nonlinear) elliptic operator. The problem will be referred to as the *stochastic elliptic boundary value problem.*
Some examples of specific problems described by the former general problem setting will be presented.

Example 2.1:
The first example is a linear second-order elliptic problem with scalar force term and solution (i.e., $\mathbf{f} = f$ and $\mathbf{w} = u$).

$$\begin{aligned} -\nabla \cdot (a(\omega, \mathbf{x})\nabla u(\omega, \mathbf{x})) &= f(\omega, \mathbf{x}) \text{ in } \Omega \times D, \mathbb{P}\text{-a.e.} \\ u(\omega, \mathbf{x}) &= g(\omega, \mathbf{x}) \text{ on } \Omega \times \partial D, \mathbb{P}\text{-a.e..} \end{aligned}$$

The linear operator is given by $\mathcal{L}(a)(u) = -\nabla \cdot (a\nabla u)$. Note that throughout the thesis the divergence operator $\nabla\cdot$ and the gradient operator ∇ are understood with respect to the spatial domain only, i.e., $\nabla = \nabla_{\mathbf{x}}, \mathbf{x} \in D$. □

Example 2.2:
In [10] the solution $\mathbf{w} = (p, \mathbf{u}) : \Omega \times \bar{D} \to \mathbb{R}^{d+1}$ of the mixed form of the linear second order elliptic problem, that is,

$$\begin{aligned} \frac{1}{a(\omega, \mathbf{x})}\mathbf{u}(\omega, \mathbf{x}) - \nabla p(\omega, \mathbf{x}) &= \mathbf{0} && \text{in } \Omega \times D, \mathbb{P}\text{-a.e.} \\ \nabla \cdot \mathbf{u}(\omega, \mathbf{x}) &= -f(\omega, \mathbf{x}) && \text{in } \Omega \times D, \mathbb{P}\text{-a.e.} \\ p(\omega, \mathbf{x}) &= g(\omega, \mathbf{x}) && \text{on } \Omega \times \partial D, \mathbb{P}\text{-a.e.,} \end{aligned}$$

is sought. The force term is given by $\mathbf{f} = (\mathbf{0}, -f) : \Omega \times \bar{D} \to \mathbb{R}^{d+1}$ and the diffusion coefficient a enters nonlinearly into the first equation. The linear operator is defined as

$$\mathcal{L}(a)(p, \mathbf{u}) = \begin{pmatrix} \frac{1}{a}\mathbf{u} - \nabla p \\ \nabla \cdot \mathbf{u} \end{pmatrix}.$$

□

Example 2.3:
This example is not relevant in the sequel of the thesis, but it is introduced to illustrate a case of a nonlinear second order elliptic problem. Let $k \in \mathbb{N}$. Then consider

$$\begin{aligned} -\nabla \cdot (a(\omega, \mathbf{x})\nabla u(\omega, \mathbf{x})) + u(\omega, \mathbf{x})|u(\omega, \mathbf{x})|^k &= f(\omega, \mathbf{x}) \text{ in } \Omega \times D, \mathbb{P}\text{-a.e.} \\ u(\omega, \mathbf{x}) &= g(\omega, \mathbf{x}) \text{ on } \Omega \times \partial D, \mathbb{P}\text{-a.e..} \end{aligned}$$

Again, it is $\mathbf{f} = f$ and $\mathbf{w} = u$ and the nonlinear operator is defined by $\mathcal{L}(a)(u) = -\nabla \cdot (a\nabla u) + u|u|^k$. □

A function space for the functions occurring in the elliptic problem formulation will be introduced. Let $(W(D), \|\mathbf{v}\|_{W(D)})$ be an arbitrary Banach space of functions $\mathbf{v} : D \to \mathbb{R}^n$ with corresponding norm $\|\mathbf{v}\|_{W(D)}$.

Definition 2.4 – The $L^q_{\mathbb{P}}(\Omega; W(D))$-space:
The space $L^q_{\mathbb{P}}(\Omega; W(D))$, $1 \leq q \leq \infty$, is the space of measurable functions $\mathbf{v} : \Omega \to W(D)$ such that the corresponding norm is finite, where

$$\|\mathbf{v}\|_{L^q_{\mathbb{P}}(\Omega;W(D))} = \begin{cases} \left(\int_\Omega \|\mathbf{v}(\omega,\cdot)\|^q_{W(D)} \mathrm{d}\mathbb{P}(\omega) \right)^{1/q} & \textit{for } 1 \leq q < \infty, \\ \operatorname*{ess\,sup}_{\omega\in\Omega} \|\mathbf{v}(\omega,\cdot)\|_{W(D)} & \textit{for } q = \infty. \end{cases}$$

□

Remark 2.5:
Whenever a vector-valued function $\mathbf{v} = (v_1, \ldots, v_n) \in W(D)$ is considered, where for instance $W(D) = L^p(D)$, the $L^p(D)$-norm of $\mathbf{v}$, $\|\mathbf{v}\|_{L^p(D)}$, can be understood as

$$\|\mathbf{v}\|_{L^p(D)} = \left(\sum_{i=1}^n \|v_i\|^p_{L^p(D)} \right)^{1/p} \quad \text{for } 1 \leq p < \infty,$$
$$\|\mathbf{v}\|_{L^\infty(D)} = \max_{i=1,\ldots,n} \|v_i\|_{L^\infty(D)} \quad \text{for } p = \infty.$$

Note that the norm $\|\mathbf{v}\|_{W(D)}$ is finite if and only if the norms $\|v_i\|_{W(D)}$ of each entry of $\mathbf{v}$ are finite. □

Weak formulation

The weak formulation of (2.1) is:
Find a function $\mathbf{w} \in L^q_{\mathbb{P}}(\Omega; W(D))$ such that for any $\mathbf{v} \in L^q_{\mathbb{P}}(\Omega; W(D))$

$$\int_\Omega \int_D \mathcal{L}\big(a(\omega,\mathbf{x})\big)\big(\mathbf{w}(\omega,\mathbf{x})\big)\mathbf{v}(\omega,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbb{P}(\omega) = \int_\Omega \int_D \mathbf{f}(\omega,\mathbf{x})\mathbf{v}(\omega,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbb{P}(\omega) \tag{2.2}$$

with additional equations for suitable boundary conditions.
The following assumption includes that the weak formulation (2.2) has a unique solution in the space $L^2_{\mathbb{P}}(\Omega; W(D))$. This fact is needed later on.

Assumption 2.6 – On the solution w of (2.1):
Let the following conditions hold true

- The solution $\mathbf{w}$ of (2.1) is in $W(D)$, i.e., $\mathbf{w}(\omega,\cdot) \in W(D)$ almost everywhere,

- there is a constant $C(\omega)$ which may depend on $\omega \in \Omega$ such that $\mathbb{P}$-a.e. the stability result

$$\|\mathbf{w}(\omega,\cdot)\|_{W(D)} \leq C(\omega)\|\mathbf{f}(\omega,\cdot)\|_{W^*(D)}$$

holds, where $W^*(D)$ is the dual space of $W(D)$ and

- $\mathbf{f} \in L^2_{\mathbb{P}}(\Omega; W^*(D))$ and the boundary condition are such that the solution $\mathbf{w}$ is unique and bounded in $L^2_{\mathbb{P}}(\Omega; W(D))$. □

2.2 Derivation of the parametric formulation

In order to obtain a parametric formulation, assumptions have to be made on both, the probability space as well as the data. During the derivation, this formulation will be defined on another probability space.

Assumption 2.7 – On the probability space:
According to [14] the random fields $a(\omega, \mathbf{x})$ and $\mathbf{f}(\omega, \mathbf{x})$ are in general not correlated. Consequently the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ might be considered as a product space of independent probability spaces $(\Omega_a, \mathcal{F}_a, \mathbb{P}_a)$ and $(\Omega_{\mathbf{f}}, \mathcal{F}_{\mathbf{f}}, \mathbb{P}_{\mathbf{f}})$, on which $a(\omega_a, \mathbf{x})$ and $\mathbf{f}(\omega_{\mathbf{f}}, \mathbf{x})$ are defined, respectively. □

In order to state the assumption on the input data, the following definition is needed.

Definition 2.8 – $\boldsymbol{\sigma}$-measurability:
Let $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$, $(A, \mathcal{A})$ and $(B, \mathcal{B})$ be measure spaces.
Let $\sigma(h) = \{h^{-1}(C) : C \in \mathcal{B}\}$ be the sigma algebra generated by the function $h : A \to B$. A function $g : \Omega \to \Omega'$ is called $\sigma(h)$-measurable if $g^{-1}(F) \in \sigma(h)$ for any $F \in \mathcal{F}'$. □

Assumption 2.9 – On the input data:
Let the input data $a(\omega_a, \mathbf{x})$, $\mathbf{f}(\omega_{\mathbf{f}}, \mathbf{x})$ fulfil the following so-called *finite-dimensional noise assumptions*

1. $a(\omega_a, \mathbf{x}) = \tilde{a}(\mathbf{y}_a(\omega_a), \mathbf{x})$ and $\tilde{a}(\mathbf{y}_a(\omega_a), \mathbf{x})$ is $\sigma(\mathbf{y}_a)$-measurable,
2. $\mathbf{f}(\omega_{\mathbf{f}}, \mathbf{x}) = \tilde{\mathbf{f}}(\mathbf{y}_{\mathbf{f}}(\omega_{\mathbf{f}}), \mathbf{x})$ and $\tilde{\mathbf{f}}(\mathbf{y}_{\mathbf{f}}(\omega_{\mathbf{f}}), \mathbf{x})$ is $\sigma(\mathbf{y}_{\mathbf{f}})$-measurable,

where the vectors

$$\mathbf{y}_a(\omega_a) = (y_{a,1}(\omega_a), \ldots, y_{a,N_a}(\omega_a)) \text{ and } \mathbf{y}_{\mathbf{f}}(\omega_{\mathbf{f}}) = (y_{\mathbf{f},1}(\omega_{\mathbf{f}}), \ldots, y_{\mathbf{f},N_{\mathbf{f}}}(\omega_{\mathbf{f}}))$$

consisting of real-valued random variables are of dimension $N_a \in \mathbb{N}$ and $N_{\mathbf{f}} \in \mathbb{N}$. Further let $\mathbf{y} = (\mathbf{y}_a, \mathbf{y}_{\mathbf{f}}) = (y_1, \ldots y_N)$ with $N = N_a + N_f$.
By abuse of notation, the functions $\tilde{a}(\mathbf{y}_a(\omega_a), \mathbf{x})$ and $\tilde{\mathbf{f}}(\mathbf{y}_{\mathbf{f}}(\omega_{\mathbf{f}}), \mathbf{x})$ will be denoted in the following by $a(\mathbf{y}_a(\omega_a), \mathbf{x})$ and $\mathbf{f}(\mathbf{y}_{\mathbf{f}}(\omega_{\mathbf{f}}), \mathbf{x})$. □

For many applications the assumption of the dependence on only a finite number N of random variables is reasonable. These N random variables are those capturing sufficient variability of the data. An example is given now.

Example 2.10 – Karhunen-Loève expansion:
In the literature and applications a truncated Karhunen-Loève expansion is often chosen to approximate the random coefficient $a(\omega, \mathbf{x})$ by a finite number of random variables, as required in Assumption 2.9.
The Karhunen-Loève expansion of $a(\omega, \mathbf{x})$ is given by

$$a(\omega, \mathbf{x}) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} b_n(\mathbf{x}) \mathbf{y}_n(\omega),$$

where $\{\mathbf{y}_n(\omega)\}_{n=1}^{\infty}$ are uncorrelated random variables with mean value 0 and variance 1, and $\{\lambda_n, b_n(\mathbf{x})\}_{n=1}^{\infty}$ is a sequence of pairs of the eigenvalues and eigenfunctions of the covariance function

$$\mathbb{COV}_a(\mathbf{x}, \mathbf{x}') = \sum_{n=1}^{\infty} \lambda_n b_n(\mathbf{x}) b_n(\mathbf{x}').$$

The sequence consists of nonnegative and decreasing eigenvalues. A truncated Karhunen-Loève expansion is given when the above sum is truncated such that it has a finite number of terms, i.e., ∞ is replaced by N:

$$a_N(\omega, \mathbf{x}) = \sum_{n=1}^{N} \sqrt{\lambda_n} b_n(\mathbf{x}) \mathbf{y}_n(\omega)$$

and

$$\mathbb{COV}_{a_N}(\mathbf{x}, \mathbf{x}') = \sum_{n=1}^{N} \lambda_n b_n(\mathbf{x}) b_n(\mathbf{x}').$$

For a more detailed description see [14]. □

Remark 2.11:
By Assumption 2.9 no truncation occurs, i.e., a truncation error has not to be considered. □

By the σ-measurability in Assumption 2.9 it can be ensured by the Doob-Dynkin lemma that the functions $a(\mathbf{y}_a(\omega_a), \mathbf{x})$ and $\mathbf{f}(\mathbf{y}_\mathbf{f}(\omega_\mathbf{f}), \mathbf{x})$ are Borel-measurable functions (and consequently integrable) of $\mathbf{y}_a$ and $\mathbf{y}_\mathbf{f}$, respectively. The integrability will be needed later on, to define integrals on the parametrized functions.

Theorem 2.12 – Doob-Dynkin lemma:
Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be measure spaces and let the function $h : \Omega \to \Omega'$ be a $\mathcal{F}$-$\mathcal{F}'$ measurable function. Then the following two properties are equivalent.

- *The function $b : \Omega \to \mathbb{R}$ is $\sigma(h)-$measurable.*
- *There exists a measurable function $g : \Omega' \to \mathbb{R}$ such that $b = g \circ h$.* $\square$

Setting now $(\Omega', \mathcal{F}') = (\mathbb{R}^{N_a}, \mathcal{B}(\mathbb{R}^{N_a}))$, $h = \mathbf{y}_a$ with $\mathbf{y}_a : \Omega \to \mathbb{R}^{N_a}$ (which is measurable as being a vector of random variables) and $b = \tilde{a} \circ \mathbf{y}_a$ with $\tilde{a} \circ \mathbf{y}_a : \Omega \to \mathbb{R}$, it follows the existence of a Borel-measurable function $g : \mathbb{R}^{N_a} \to \mathbb{R}$ such that $\tilde{a} \circ \mathbf{y}_a = g \circ \mathbf{y}_a$. By construction, it holds necessarily $g = \tilde{a}$ and consequently $\tilde{a}(\mathbf{y}_a(\omega_a), \mathbf{x}) = a(\mathbf{y}_a(\omega_a), \mathbf{x})$ is Borel-measurable with respect to $\mathbf{y}_a$. The same argument can be used for the input data $\mathbf{f}$ by applying the above theorem to each entry in the vector $\mathbf{f}$.
Still, the aim is to parametrize the stochastic elliptic problem. By the above defined vector of independent random variables

$$\mathbf{y} : \Omega \to \Gamma \subset \mathbb{R}^N, N \in \mathbb{N} \quad \text{with} \quad y_n : \Omega \to \Gamma_n \subset \mathbb{R}$$

the parameter domain Γ is given by the product of the one-dimensional images of each random variable

$$\Gamma = \Gamma_1 \times \cdots \times \Gamma_N = \prod_{n=1}^{N} \Gamma_N.$$

Further, the following assumption will be needed.

Assumption 2.13 – Joint probability density function:
Let the image measure of $\mathbf{y}$ be given by a joint density function $\rho(\mathbf{y})$ with respect to the Lebesgue measure $\mathrm{d}\mathbf{y}$

$$\rho(\mathbf{y}) : \Gamma \to \mathbb{R}_+ \text{ with } \rho(\mathbf{y}) \in L^\infty(\Gamma).$$

For example, this assumption is fulfilled if the random variables are independent or if $\mathbf{y}(\omega)$ is absolutely continuous with respect to the Lebesgue measure. As needed later for the collocation method (see Section 4.1), the density ρ is assumed to factorize, i.e.,

$$\rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n).$$

$\square$

Remark 2.14:
Note that here the density ρ is assumed to factorize. In the case of a non factorizing density ρ it is described in [1, p. 1011] that an auxiliary probability density function $\hat{\rho} : \Gamma \to \mathbb{R}_+$ is introduced which factorizes and is such that $\|\frac{\rho}{\hat{\rho}}\|_{L^\infty(\Gamma)} < \infty$. □

Now, it has to be described how the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ can be transferred to a probability space on the parameter domain Γ, denoted in the following by $(\Gamma, \mathcal{B}(\Gamma), \rho(\mathbf{y})\,\mathrm{d}\mathbf{y})$.
Given the three assumptions stated above, it is possible to map the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to the space $(\Gamma, \mathcal{B}(\Gamma), \rho(\mathbf{y})\,\mathrm{d}\mathbf{y})$ by using the concept of the image measure.

Definition 2.15 – Image measure:
Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, let $(\Omega', \mathcal{F}')$ be a measurable space, and let $h : \Omega \to \Omega'$ be a $\mathcal{F}$-$\mathcal{F}'$ measurable function. Then

$$\nu(F_2) = \mu(h^{-1}(F_2)) \text{ with } F_2 \in \mathcal{F}'$$

defines a measure (called image measure or pushforward measure) $\nu = \mu \circ h^{-1}$ on $(\Omega', \mathcal{F}')$. □

The function $\mathbf{y}$ is the measurable function of the above definition from the measure space $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\Gamma, \mathcal{B}(\Gamma))$. Again, it is measurable because the y_i are random variables defined on Ω with image in $\mathbb{R}$.
Further, the measure $\mathbb{P} \circ \mathbf{y}^{-1}$ is an image measure on $(\Gamma, \mathcal{B}(\Gamma))$, where $\mathcal{B}(\Gamma) \subset \mathcal{B}(\mathbb{R}^N)$. In Assumption 2.13 this image measure was assumed to be given by $\rho(\mathbf{y})$ with respect to the Lebesgue measure, i.e., $\mathbb{P} \circ \mathbf{y}^{-1} = \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}$. Thus, the space $(\Gamma, \mathcal{B}(\Gamma), \rho(\mathbf{y})\,\mathrm{d}\mathbf{y})$ is obtained.

The last step to obtain the parametric formulation is the following: Show that under certain conditions the solution $\mathbf{w}(\omega, \mathbf{x})$ can be represented by a finite number of random variables too, i.e., $\mathbf{w}(\omega, \mathbf{x}) = \tilde{\mathbf{w}}(\mathbf{y}(\omega), \mathbf{x})$. As before, $\mathbf{w}(\mathbf{y}(\omega), \mathbf{x})$ will be written instead of $\tilde{\mathbf{w}}(\mathbf{y}(\omega), \mathbf{x})$.
With this, problem (2.1) would be given by:
Find a solution $\mathbf{w} : \Gamma \times \bar{D} \to \mathbb{R}^n$ such that almost everywhere in Γ the following holds

$$\mathcal{L}(a(\mathbf{y}, \mathbf{x}))(\mathbf{w}(\mathbf{y}, \mathbf{x})) = \mathbf{f}(\mathbf{y}, \mathbf{x}) \text{ in } D, \tag{2.3}$$

with additional equations for suitable boundary conditions.

In order to verify $\mathbf{w}(\omega, \mathbf{x}) = \mathbf{w}(\mathbf{y}(\omega), \mathbf{x})$, the weak forms of the stochastic and parametric problem as well as a result from measure theory on the image measure given in Theorem 2.20 are of importance .

Before giving the weak formulation and applying the theorem, function spaces have to be defined. Define the following Bochner-Lebesgue-spaces, where, as before, $\mathbf{v} : D \to \mathbb{R}^n$ with corresponding norm $\|\mathbf{v}\|_{W(D)}$. The following definition is required although it can be shown by Theorem 2.20 to be the same as Definition 2.4.

Definition 2.16 – The $L^q_\rho(\Gamma; W(D))$-space:
The space $L^q_\rho(\Gamma; W(D))$, $1 \leq q \leq \infty$, is the space of measurable functions $\mathbf{v} : \Gamma \to W(D)$ such that the corresponding norm is finite, where

$$\|\mathbf{v}\|_{L^q_\rho(\Gamma;W(D))} = \begin{cases} \left(\int_\Gamma \|\mathbf{v}\|^q_{W(D)} \rho(\mathbf{y}) \, \mathrm{d}\mathbf{y} \right)^{1/q} & \textit{for } 1 \leq q < \infty, \\ \operatorname*{ess\,sup}_{\mathbf{y} \in \Gamma} \|\mathbf{v}\|_{W(D)} & \textit{for } q = \infty. \end{cases} \qquad \square$$

Remark 2.17:
Note that in the case $q = 2$ (which will be the relevant case in this thesis) and if $W(D)$ is a Hilbert space, the spaces $L^2_\rho(\Gamma; W(D))$ and $L^2_\rho(\Gamma) \otimes W(D)$ are isomorphic. $\square$

Multiplying the equation (2.3) by a function $\mathbf{v} \in L^q_\rho(\Gamma; W(D))$ the weak formulation is obtained.

Weak formulation of (2.3)

Find a function $\mathbf{w} \in L^q_\rho(\Gamma; W(D))$ such that for any $\mathbf{v} \in L^q_\rho(\Gamma; W(D))$ it holds

$$\int_{\Gamma \times D} \mathcal{L}\big(a(\mathbf{y}, \mathbf{x})\big)\big(\mathbf{w}(\mathbf{y}, \mathbf{x})\big)\mathbf{v}(\mathbf{y}, \mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \int_{\Gamma \times D} \mathbf{f}(\mathbf{y}, \mathbf{x})\mathbf{v}(\mathbf{y}, \mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \tag{2.4}$$

with additional equations for suitable boundary conditions.

Assumption 2.18 – Uniqueness of w on $\Gamma \times D$:
Let the same assumptions as in Assumption 2.6 hold true on the solution $\mathbf{w}$ to (2.4), where Ω is replaced by Γ and ω by $\mathbf{y}$. In particular it is assumed that the solution is unique. Further, assume that $\mathbf{w}(\mathbf{y}(\omega), \mathbf{x})$ is $\sigma(\mathbf{y})$-measurable.$\square$

Remark 2.19:
The assumption that $\mathbf{w}(\mathbf{y}(\omega), \mathbf{x})$ is $\sigma(\mathbf{y})$-measurable ensures – analogue to the data – via the Doob-Dynkin lemma the Borel-measurability and the integrability of $\mathbf{w}(\mathbf{y}(\omega), \mathbf{x})$. Consequently it is possible to set up the weak formulation. $\square$

With the following Theorem 2.20 (see for example in [9, p. 191]) two things can be clarified:

- The weak formulation (2.2) of (2.1) depending on functions defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equals the weak formulation (2.4) of the parametrized stochastic problem (2.3) of functions defined on $(\Gamma, \mathcal{B}(\Gamma), \rho(\mathbf{y})\,\mathrm{d}\mathbf{y})$.
- The solution $\mathbf{w}(\omega, \mathbf{x})$ can also be represented by a finite number of random variables, i.e., $\mathbf{w}(\omega, \mathbf{x}) = \mathbf{w}(\mathbf{y}(\omega), \mathbf{x})$.

Theorem 2.20:
Let $\mu \circ h^{-1}$ be the image measure as in Definition 2.15. Let $\mathbf{g} : \Omega' \to \mathbb{R}^n$ be a $\mathcal{F}'$– $\mathcal{B}(\mathbb{R}^n)$ measurable function. Then $\mu \circ h^{-1}$ almost everywhere the following equality holds (where $F_2 \subset \mathcal{F}'$)

$$\int_{h^{-1}(F_2)} \mathbf{g} \circ h \ \mathrm{d}\mu = \int_{F_2} \mathbf{g} \ \mathrm{d}(\mu \circ h^{-1}).$$

$\square$

Making the ansatz

$$\mathbf{w}(\omega, \mathbf{x}) = \mathbf{w}(\mathbf{y}(\omega), \mathbf{x}) \tag{2.5}$$

the integrals in (2.2) can be transformed to the integrals in (2.4) by Theorem 2.20.
For this purpose, the function $\mathbf{g}$ which appears in the theorem is chosen to equal $\mathbf{f}(\mathbf{y}, \mathbf{x})$ or $\mathcal{L}\big(a(\mathbf{y}, \mathbf{x})\big)\big(\mathbf{w}\big(\mathbf{y}, \mathbf{x})\big)$. The transformation will be given for $\mathbf{f}$ and can be performed equivalently for $\mathcal{L}(a)(\mathbf{w})$.
As above, it is $h = \mathbf{y} : \Omega \to \Gamma$, $\mathrm{d}\mu = \mathrm{d}\mathbb{P}(\omega)$, $\mathrm{d}(\mu \circ h^{-1}) = \rho(\mathbf{y})\mathrm{d}\mathbf{y}$. Then it follows

$$\begin{aligned}
&\int_D \int_\Omega \mathbf{f}(\omega, \mathbf{x})\mathbf{v}(\omega, \mathbf{x}) \,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}\mathbf{x} \\
=& \int_D \int_{\mathbf{y}^{-1}(\Gamma)} \mathbf{f}(\mathbf{y}(\omega), \mathbf{x})\mathbf{v}(\mathbf{y}(\omega), \mathbf{x}) \,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}\mathbf{x} \\
=& \int_D \int_\Gamma \mathbf{f}(\mathbf{y}, \mathbf{x})\mathbf{v}(\mathbf{y}, \mathbf{x})\rho(\mathbf{y}) \,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}.
\end{aligned}$$

The first equality holds by assumption, the second by application of Theorem 2.20.
The solutions coincide and thus the ansatz (2.5) is confirmed because 1. the integrals of both weak formulations are equal to each other (note that by the Fubini lemma, the order of integration can be interchanged) and 2. $\mathbf{w}(\omega, \mathbf{x})$ is the unique solution of (2.2) and $\mathbf{w}(\mathbf{y}, \mathbf{x})$ is the unique solution of (2.4). Therefore, the parametric formulation as given in (2.3) holds.

3 Diffusion problem and its mixed form – Existence, uniqueness and analytic extension

In this section assumptions on the data in the diffusion problem and its mixed form are formulated. With these assumptions the respective problem admits a unique solution. The mixed form of the diffusion problem is considered because in many applications the flux of the solution is the quantity of interest (see [11], p. 2), often more than the solution itself.
The analytic extension of the solution will be crucial to derive approximation error estimates. The approximations considered further in this thesis (see Section 4) are the tensor product approximation and the sparse grid Smolyak approximation via the collocation method (see Section 4.1). The same grid constructions are considered for numerical quadrature (see Section 4.2). In both approaches the best-approximation error has to be bounded and the analytic extension is needed.
Depending on the grid choice (tensor product or sparse grid) different results on the extensibility of the solution to the complex plane are required. For the first grid, the solution has to be extended in one coordinate direction, for the second in all coordinate directions simultaneously. The results on the analytic extensibility are of relevance for the error estimates in Section 4.4.
The last part of this section states additional regularity assumptions which are of relevance in the multilevel method examined in Section 5.

3.1 The diffusion problem and its mixed form

Consider the stochastic diffusion problem in standard form, which was already given in Example 2.1,

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

and its parametric equivalent (the parametric diffusion problem)

$$\begin{aligned} -\nabla \cdot (a(\mathbf{y}, \mathbf{x})\nabla u(\mathbf{y}, \mathbf{x})) &= f(\mathbf{y}, \mathbf{x}) \text{ in } \Gamma \times D, \quad \rho(\mathbf{y})\mathrm{d}\mathbf{y}, \\ u(\mathbf{y}, \mathbf{x}) &= g(\mathbf{y}, \mathbf{x}) \text{ on } \Gamma \times \partial D, \ \rho(\mathbf{y})\mathrm{d}\mathbf{y}. \end{aligned} \tag{3.1}$$

Remark 3.1:
From now on only the parametric diffusion problem is considered. According to the parametric diffusion problem, the same properties can be derived for the stochastic diffusion problem. □

The mixed form of the diffusion problem was given in Example 2.2. It is equivalent with the diffusion problem as the following reasoning shows. Hence, it seems to be natural to consider in the analysis of this section both, the diffusion problem and its mixed form, and to identify similarities. Consider the parametrized form of the diffusion problem's mixed form

$$\begin{aligned}
\frac{1}{a(\mathbf{y},\mathbf{x})}\mathbf{u}(\mathbf{y},\mathbf{x}) - \nabla p(\mathbf{y},\mathbf{x}) &= 0 && \text{in } \Gamma \times D, \quad \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\text{-a.e.},\\
\nabla \cdot \mathbf{u}(\mathbf{y},\mathbf{x}) &= -f(\mathbf{y},\mathbf{x}) && \text{in } \Gamma \times D, \quad \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\text{-a.e.},\\
p(\mathbf{y},\mathbf{x}) &= g(\mathbf{y},\mathbf{x}) && \text{on } \Gamma \times \partial D, \ \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\text{-a.e.}.
\end{aligned} \tag{3.2}$$

The first line yields $\frac{1}{a(\mathbf{y},\mathbf{x})}\mathbf{u}(\mathbf{y},\mathbf{x}) = \nabla p(\mathbf{y},\mathbf{x})$. Substituting $a(\mathbf{y},\mathbf{x})\nabla p(\mathbf{y},\mathbf{x})$ for $\mathbf{u}(\mathbf{y},\mathbf{x})$ in the second equation and multiplying by -1 leads to

$$\begin{aligned}
-\nabla \cdot (a(\mathbf{y},\mathbf{x})\nabla p(\mathbf{y},\mathbf{x})) &= f(\mathbf{y},\mathbf{x}) && \text{in } \Gamma \times D, \quad \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\text{-a.e.},\\
p(\mathbf{y},\mathbf{x}) &= g(\mathbf{y},\mathbf{x}) && \text{on } \Gamma \times \partial D, \ \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\text{-a.e.},
\end{aligned}$$

which equals the diffusion problem described in (3.1), with $p = u$.

3.2 Different forms of the diffusion coefficient

Two different forms of the diffusion coefficient $a(\omega, \mathbf{x})$ (or $a(\mathbf{y}, \mathbf{x})$ in the parametrized form) will be considered: Either the diffusion coefficient is uniformly bounded, i.e., there are $0 < a_{min}$ and $a_{max} < \infty$ such that

$$0 < a_{min} \leq a(\omega, \mathbf{x}) \leq a_{max} < \infty, \tag{3.3}$$

or the diffusion coefficient is only bounded by random variables $0 < a_{min}(\omega)$ and $a_{max}(\omega) < \infty$ such that

$$0 < a_{min}(\omega) = \operatorname*{ess\,inf}_{\mathbf{x}\in D} a(\omega, \mathbf{x}) \leq a(\omega, \mathbf{x}) \leq \operatorname*{ess\,sup}_{\mathbf{x}\in D} a(\omega, \mathbf{x}) = a_{max}(\omega) < \infty. \tag{3.4}$$

Bounded by random variables will in the following be referred to as *D-bounded* as the diffusion coefficient $a(\omega, \mathbf{x})$ is bounded in the spatial variable.

Remark 3.2:
Strictly speaking, a uniformly bounded is also a D-bounded diffusion coefficient. Nevertheless, both cases are considered on their own as more assumptions on the data need to be stated for a D-bounded diffusion coefficient. □

Example 3.3 – Uniformly bounded diffusion coefficient:
Assume that $a > a_{min} > 0$. Consider the Karhunen-Loève expansion (see Example 2.10) for $\log(a - a_{min})$, where $\mathbb{E}[\log\big(a(\omega, \mathbf{x})\big)] = 0$. It holds

$$\log(a(\omega, \mathbf{x}) - a_{min}) = \sum_{n=1}^{N} \sqrt{\lambda_n} b_n(\mathbf{x}) y_n(\omega),$$

and consequently

$$a(\omega, \mathbf{x}) = a_{min} + \exp\left(\sum_{n=1}^{N} \sqrt{\lambda_n} b_n(\mathbf{x}) y_n(\omega)\right).$$

Due to the last expression $a(\omega, \mathbf{x})$ is uniformly bounded away from zero for any choice of random variables $\{y_n(\omega)\}_{n=1}^{N}$. It is uniformly bounded if for instance the y_n are uniformly distributed in $[-1; 1]$ and $b_n \in L^\infty(D)$ (see Definition B.1). For standard normally distributed y_n, however, the diffusion coefficient cannot be uniformly bounded from above. $\square$

Remark 3.4 – Lognormal diffusion coefficient:
The coefficient is often lognormally distributed in applications, i.e., its logarithm is normally distributed, $\log\big(a(\omega, \mathbf{x})\big) \sim \mathcal{N}(\mu, \sigma^2)$. This is a special case of a D-bounded diffusion coefficient which becomes clear by the following example. $\square$

Example 3.5 – Lognormal diffusion coefficient:
Assume that $\log\big(a(\omega, \mathbf{x})\big) \sim \mathcal{N}(0, \sigma^2)$. Consider the following lognormal distributed diffusion coefficient, obtained through the Karhunen-Loève expansion (see Example 2.10),

$$a(\omega, \mathbf{x}) = \exp\left(\sum_{n=1}^{N} \sqrt{\lambda_n} b_n(\mathbf{x}) y_n(\omega)\right),$$

where the $y_n \sim \mathcal{N}(0, 1)$, $n \in \{1, \dots, N\}$, are independent and identically distributed. Since the y_n are in $(-\infty, \infty)$, the exponent cannot be uniformly bounded by a constant bigger than zero. $\square$

The last example shows that for a lognormal distribution it holds no longer $a(\omega, \mathbf{x}) \geq a_{min} > 0$. Rather, the diffusion coefficient is in that case assumed to fulfil the condition on a D-bounded diffusion coefficient (3.4).

Remark 3.6:
Consider again the previous example. Let the functions $b_n \in L^\infty(D)$. Then (after parametrizing), the bounds of (3.4) are satisfied because

$$0 < \exp\Big(- \sum_{n=1}^{N} \sqrt{\lambda_n} \|b_n\|_{L^\infty(D)} |y_n| \Big) \leq a_{min}(\mathbf{y}) = \operatorname*{ess\,inf}_{\mathbf{x}\in D} a(\mathbf{y}, \mathbf{x})$$

and

$$\operatorname*{ess\,sup}_{\mathbf{x}\in D} a(\mathbf{y}, \mathbf{x}) = \|a(\mathbf{y}, \cdot)\|_{L^\infty(D)} \leq \exp\Big(\sum_{n=1}^{N} \sqrt{\lambda_n} \|b_n\|_{L^\infty(D)} |y_n| \Big) < \infty. \quad \square$$

3.2.1 Different forms of the parameter domain

The form of the diffusion coefficient as well as the choice of the distribution of the random variables give implications on the parameter space Γ. The cases whether Γ is bounded or unbounded have to be distinguished.
For a lognormal diffusion coefficient in the form as in Example 3.5 the parameter space Γ_n equals $\mathbb{R}$, i.e., it is unbounded and so is the whole parameter space Γ. This is because the random variables $\mathbf{y}_n$, $n \in \{1, \ldots, N\}$, are $\mathcal{N}(0, 1)$ distributed, i.e., $\rho_n(y_n) = \frac{1}{2\pi} \exp(\frac{-y_n^2}{2})$. For other forms of D-bounded diffusion coefficients a bounded Γ might be possible.
In the case of a uniformly bounded diffusion coefficient several choices are possible as well, depending on the form of the approximation of the random field. Taking the approximation as in Example 3.3, for a uniform distribution in some closed interval of the random variables y_n the parameter space Γ_n is bounded.
Hence it is of interest to consider bounded and unbounded parameter spaces.
In the case of Γ bounded it is assumed (without loss of generality), that $\Gamma = [-1, 1]^N$ because every interval $[a, b]$ can be transformed to $[-1, 1]$ by the mapping $t \mapsto 2(t - \frac{a+b}{2})/(b - a)$.
In order to treat bounded and unbounded parameter domains, introduce – as described in [1] – a weight

$$\sigma : \Gamma \to \mathbb{R}_+$$

with $\sigma(\mathbf{y}) = \prod_{n=1}^{N} \sigma_n(y_n) \leq 1$, and $\sigma_n(y_n) = 1$ if Γ_n is bounded.
Based on this weight, the following function space is defined.

Definition 3.7:
The space $C^0_\sigma(\Gamma; W(D)) =$

$$\left\{\mathbf{v} : \Gamma \to W(D) : \mathbf{v} \textit{ continuous in } \mathbf{y}, \max_{\mathbf{y}\in\Gamma} \|\sigma(\mathbf{y})\mathbf{v}(\mathbf{y})\|_{W(D)} < \infty\right\}$$

$$\textit{with norm } \|\mathbf{v}\|_{C^0_\sigma(\Gamma;W(D))} = \max_{\mathbf{y}\in\Gamma} \|\sigma(\mathbf{y})\mathbf{v}(\mathbf{y})\|_{W(D)}$$

consists of continuous functions $\mathbf{v} : \Gamma \to W(D)$ *whose norm* $\|\cdot\|_{C^0_\sigma(\Gamma;W(D))}$ *is finite.* □

Remark 3.8:
Note that the functions in the above definition are understood as functions from Γ to $W(D)$. Subsequently – if no ambiguity arises – sometimes $v(\mathbf{y})$ will be written instead of $v(\mathbf{y}, \cdot)$, and the function v will be understood as a function from Γ to $W(D)$. □

Remark 3.9:
If Γ_n is bounded, it is $\sigma_n(y_n) = 1$. Consequently, if the whole parameter space Γ is bounded, it holds $\sigma(\mathbf{y}) = 1$, and the just defined space $C^0_\sigma(\Gamma; W(D))$ equals the space $C^0(\Gamma; W(D))$ of continuous functions with bounded maximum-norm. □

3.3 Existence and uniqueness results

While existence, uniqueness, and continuity results on the parameter space Γ in $\mathbb{R}^N$ are sufficient for deriving an analytic extension in one direction, results on parameter spaces extended to $\mathbb{C}^N$ are required to get an analytic extension in all variables simultaneously. An analytic extension in all variables simultaneously is necessary for approximations or quadratures based on Smolyak sparse grids and quasi-optimal sparse grids (see Section 4.1.3), while an analytic extension in one direction suffices on tensor-product grids. Thus, the diffusion problem and its mixed form are examined on the more general space $\mathbb{C}^N$. Conditions on the data and spaces are stated in order to obtain existence and uniqueness of the weak solution. Also, to motivate the approximation by global polynomials in the parameter domain and to derive analyticity, a continuity result for the solution is shown.
This subsection will be concluded by Section 3.3.3, where the connection to the problems defined in the real plane will be pointed out. Also, it will be shown that the solution exists in $L^2_\rho(\Gamma; W(D))$ with $\Gamma \subset \mathbb{R}^N$ under certain conditions.
Before examining the diffusion problem and its mixed form, the data are assumed to be extendible to the complex plane.

Let $\mathbf{z} = (z_1, \dots, z_N) \in \mathbb{C}^N$ be a complex valued variable with $\mathrm{Re}(\mathbf{z}) = \mathbf{y}$. The following assumption on the data has to be fulfilled.

Assumption 3.10 – Extensions of the data to complex plane:
Let $a(\cdot, \mathbf{x})$ and $f(\cdot, \mathbf{x})$ have extensions to the complex plane $\mathbb{C}^N$, i.e., for all $\mathbf{z} \in \mathbb{C}^N$ the data $a(\mathbf{z}, \mathbf{x})$ and $f(\mathbf{z}, \mathbf{x})$ are defined taking values in $\mathbb{C}$. □

Remark 3.11 – On the procedure:
For the uniformly bounded diffusion coefficient and the diffusion coefficient bounded by random variables the procedure is as follows.

1. Parameter spaces being subsets of $\mathbb{C}^N$ are defined on which the subsequent analysis will be performed.
2. For the diffusion problem
   a. existence and uniqueness of a weak solution with corresponding stability condition under certain assumptions on the data and spaces are shown, and
   b. a continuous solution is obtained under continuity assumptions on the data.
3. For the diffusion problem's mixed form the previous steps a. and b. are carried out too. □

The cases of a uniformly bounded and a D-bounded diffusion coefficient are now distinguished, and similarities of both problem formulations become clear and are pointed out.

3.3.1 Uniformly bounded diffusion coefficient in $\mathbb{C}^N$

Throughout the entire subsection the following assumption should hold.

Assumption 3.12 – Diffusion coefficient uniformly bounded:
Let the diffusion coefficient $a(\mathbf{y}, \cdot)$ be uniformly bounded, i.e., there exist $a_{min}, a_{max} \in (0, \infty)$ such that $\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}$-a.e. it holds

$$a_{min} \leq a(\mathbf{y}, \mathbf{x}) \leq a_{max} \text{ for all } \mathbf{x} \in \bar{D}.$$

□

The spaces needed for the analysis will be defined. This is step 1 of the procedure described in Remark 3.11.

Definition 3.13 – The set Σ_U:
Let $0 < \bar{a}_{min} \le a_{min}$ and $a_{max} \le \bar{a}_{max} < \infty$ with a_{min}, a_{max} as in Assumption 3.12. Define the set Σ_U – the index $_U$ indicates the uniformly bounded case – as

$$\Sigma_U = \left\{ \mathbf{z} \in \mathbb{C}^N : \bar{a}_{min} \le \operatorname{Re}(a(\mathbf{z}, \mathbf{x})) \le |a(\mathbf{z}, \mathbf{x})| \le \bar{a}_{max} \quad \forall \mathbf{x} \in \bar{D} \right\}. \qquad \square$$

Remark 3.14 – $\boldsymbol{\Gamma \subset \Sigma_U}$:
The previous assumption implies $\Gamma \subset \Sigma_U$. Since for every $\mathbf{y} \in \Gamma$ it is $\mathbf{y} \in \mathbb{R}^N \subset \mathbb{C}^N$, $\operatorname{Re}(a(\mathbf{y}, \mathbf{x})) = a(\mathbf{y}, \mathbf{x})$ and $|a(\mathbf{y}, \mathbf{x})| = a(\mathbf{y}, \mathbf{x})$, it follows

$$\bar{a}_{min} \le a_{min} \le \operatorname{Re}(a(\mathbf{y}, \mathbf{x})) = a(\mathbf{y}, \mathbf{x}) \le a_{max} \le \bar{a}_{max} \quad \forall \mathbf{x} \in \bar{D},$$

and thus $\mathbf{y} \in \Sigma_U$. $\square$

The space of continuous functions of Γ to the Banach space $W(D)$ is extended to functions from Σ_U to $W(D)$.

Definition 3.15 – The space $\boldsymbol{C^0_\sigma(\Sigma_U; W(D))}$:
For a weight function $\sigma : \mathbb{R}^N \to \mathbb{R}_+$, let

$$C^0_\sigma(\Sigma_U; W(D)) = \{\mathbf{v} : \Sigma_U \to W(D) : \mathbf{v} \textit{ continuous in } \mathbf{z}, \\ \max_{\mathbf{z} \in \Sigma_U} \|\sigma(\operatorname{Re} \mathbf{z}) \mathbf{v}(\mathbf{z})\|_{W(D)} < \infty\}$$

be the space of continuous functions from Σ_U to the Banach space $W(D)$ with norm

$$\|\mathbf{v}\|_{C^0_\sigma(\Sigma_U; W(D))} = \max_{\mathbf{z} \in \Sigma_U} \|\sigma(\operatorname{Re} \mathbf{z}) \mathbf{v}(\mathbf{z})\|_{W(D)}. \qquad \square$$

Subsequently, the steps 2 and 3 – each with the parts a and b – as stated in Remark 3.11 are carried out, for the diffusion problem and its mixed form.

3.3.1.1 Diffusion problem

As mentioned before, results on subspaces of $\mathbb{C}^N$ are of interest. Hence, the values in the stochastic variable are taken in $\mathbb{C}^N$.

Assumption 3.16 – Spaces of the diffusion problem:
Let $W(D) = H^1_0(D)$ and $q = 2$. Then $L^q_\rho(\Gamma; W(D)) = L^2_\rho(\Gamma; H^1_0(D))$ with corresponding norm $\|v\|_{L^2_\rho(\Gamma; H^1_0(D))}$.
The assumption means that the solution equals zero on the boundary, i.e., $g = 0$. The space $H^1_0(D)$ with its norm is defined in Remark B.4. $\square$

Remark 3.17:
It is sufficient to consider only the diffusion problem with a homogeneous Dirichlet boundary condition. A problem with a non-homogeneous Dirichlet boundary condition can be transformed into a problem with homogeneous Dirichlet boundary condition. Let

$$\begin{aligned} \nabla \cdot \big(a(\mathbf{z},\mathbf{x})\nabla w(\mathbf{z},\mathbf{x})\big) &= \tilde{f}(\mathbf{z},\mathbf{x}) \quad \text{on } \Gamma \times D, \\ w(\mathbf{z},\mathbf{x}) &= g(\mathbf{z},\mathbf{x}) \quad \text{on } \Gamma \times \partial D \end{aligned}$$

be the original problem. Let v be an arbitrary function on $\Gamma \times \bar{D}$ such that $v|_{\partial D} = g$. Then, inserting $w = v + u$, the following equations

$$\begin{aligned} \nabla \cdot \big(a(\mathbf{z},\mathbf{x})\nabla (v+u)(\mathbf{z},\mathbf{x})\big) &= \tilde{f}(\mathbf{z},\mathbf{x}) \quad \text{on } \Gamma \times D, \\ (v+u)(\mathbf{z},\mathbf{x}) &= g(\mathbf{z},\mathbf{x}) \quad \text{on } \Gamma \times \partial D \end{aligned}$$

are obtained, which are due to the linearity of the divergence and the gradient operator equal to

$$\begin{aligned} \nabla \cdot \big(a(\mathbf{z},\mathbf{x})\nabla u(\mathbf{z},\mathbf{x})\big) &= \overbrace{\tilde{f}(\mathbf{z},\mathbf{x}) - \nabla \cdot \big(a(\mathbf{z},\mathbf{x})\nabla v(\mathbf{z},\mathbf{x})\big)}^{f(\mathbf{z},\mathbf{x})} \quad && \text{on } \Gamma \times D, \\ u(\mathbf{z},\mathbf{x}) &= g(\mathbf{z},\mathbf{x}) - v(\mathbf{z},\mathbf{x}) = 0 && \text{on } \Gamma \times \partial D. \end{aligned}$$

With the solution to

$$\begin{aligned} \nabla \cdot \big(a(\mathbf{z},\mathbf{x})\nabla u(\mathbf{z},\mathbf{x})\big) &= f(\mathbf{z},\mathbf{x}) \quad && \text{on } \Gamma \times D, \\ u(\mathbf{z},\mathbf{x}) &= 0 && \text{on } \Gamma \times \partial D \end{aligned}$$

at hand, the solution of the original problem is given by $w = v + u$. □

The weak formulation of the diffusion problem in $\mathbb{C}^N$ with $\mathbf{z} \in \mathbb{C}^N$ reads as follows: Find a function $u(\mathbf{z},\cdot) \in H_0^1(D)$ such that $\rho(\mathbf{z})\,\mathrm{d}\mathbf{z}$-a.e.

$$\int_D a(\mathbf{z},\mathbf{x})\nabla u(\mathbf{z},\mathbf{x})\nabla v(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_D f(\mathbf{z},\mathbf{x})v(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad \forall v \in H_0^1(D). \tag{3.5}$$

Remark 3.18:
Note that the form (3.5) is equivalent to (2.4). The equivalence will be discussed in Section 4.1 in more detail. □

In order to obtain existence and uniqueness of the weak solution for each $\mathbf{z} \in \Sigma_U$ (step $2a$) let the following assumption on the data f hold.

Assumption 3.19 – Force term of the diffusion problem:
Let $\rho : \Sigma_U \to \mathbb{R}_+$ be a probabilistic density function. Then assume that the force term $f(\mathbf{z}, \cdot)$ with $\mathbf{z} \in \Sigma_U$ is $\rho(\mathbf{z})\,\mathrm{d}\mathbf{z}$-a.e. in $L^2(D)$. □

Remark 3.20:
No assumption on the diffusion coefficient in $\mathbb{C}^N$ is given as it is indirectly done by the definition of Σ_U. □

Theorem 3.21 – Existence of a unique solution of (3.5):
With Assumption 3.10, i.e., the assumption on the extension of the data to the complex plane, and the Assumptions 3.12, 3.16, 3.19 on the data and the spaces, for every $\mathbf{z} \in \Sigma_U$ *there exists a unique solution* $u(\mathbf{z}, \cdot)$ *in* $H_0^1(D)$ *of (3.5) satisfying* $\rho(\mathbf{z})\,\mathrm{d}\mathbf{z}$ *- a.e.*

$$\|u(\mathbf{z}, \cdot)\|_{H_0^1(D)} \leq \frac{C_P}{\bar{a}_{min}} \|f(\mathbf{z}, \cdot)\|_{L^2(D)}, \tag{3.6}$$

where C_P *is the Poincaré constant (see Lemma B.9).* □

Proof: For the proof a complex valued version of the Lemma of Lax-Milgram, see Lemma B.8, on the bilinear form $B(u, v) = \int_D a(\mathbf{z}, \mathbf{x}) \nabla u(\mathbf{z}, \mathbf{x}) \cdot \nabla v(\mathbf{x})\,\mathrm{d}\mathbf{x}$ has to be applied. The inequality follows from the Poincaré inequality given in Lemma B.9. ■

Remark 3.22:
A slightly different formulation was given in [1], where the coefficient was only assumed to be bounded from below (and only the real case was considered). In [1], similar results are given for a space equipped with some energy norm. □

Remark 3.23:
Assumption 3.19 on f to be in $L^2(D)$ is more specific than the assumption on f with respect to the spatial variable given in the general Assumption 2.6 because $L^2(D) \subset H^{-1}(D)$. $H^{-1}(D)$ is the dual space of $W(D) = H_0^1(D)$. For the norms it holds

$$\|f\|_{H^{-1}(D)} \leq C_P \|f\|_{L^2(D)},$$

where C_P, again, is the Poincaré constant of D.
Later, in the lemmas of Section 3.3.3 on the results in $\mathbb{R}^N$, all points of Assumption 2.6 are fulfilled on the smaller space $L^2(D)$. □

The case of the diffusion problem with uniformly bounded diffusion coefficient is concluded by step $2b$ from Remark 3.11.

Assumption 3.24 – Continuous data:
Let $f \in C^0_\sigma(\Sigma_U; L^2(D))$ and $a \in C^0(\Sigma_U; L^\infty(D))$. □

The following result is needed when stating the existence of an analytic extension. It also ensures that the approximation of the solution by global polynomials makes sense.

Lemma 3.25 – Continuous solution of (3.5):
With the assumptions of Theorem 3.21 and Assumption 3.24 the solution of (3.5) satisfies $u \in C^0_\sigma(\Sigma_U; H^1_0(D))$. □

Proof: The continuity of the solution $u(\mathbf{z})$ with respect to $\mathbf{z} \in \Sigma_U$ follows by the continuity of the data. The proof is similar to the proof with deterministic force term given in [7, pp. 20–21]. Due to the continuity of the data, for each ϵ there exists δ such that for all $\mathbf{z}, \tilde{\mathbf{z}} \in \Sigma_U$ with $|\mathbf{z} - \tilde{\mathbf{z}}| \leq \delta$ it holds

$$\|a(\mathbf{z}) - a(\tilde{\mathbf{z}})\|_{L^\infty(D)} \leq C\frac{\epsilon}{2} \quad \text{and} \tag{3.7}$$

$$\|f(\mathbf{z}) - f(\tilde{\mathbf{z}})\|_{L^2(D)} \leq \frac{\bar{a}_{min}}{C_P}\frac{\epsilon}{2}, \tag{3.8}$$

where the choice of the constant $C = \frac{\bar{a}^2_{min}}{C_P\|f(\tilde{\mathbf{z}})\|_{L^2(D)}}$ becomes clear subsequently. Denote by $u = u(\mathbf{z})$ the solution of the diffusion problem (3.5) with data $a = a(\mathbf{z})$ and $f = f(\mathbf{z})$ and by $\tilde{u} = u(\tilde{\mathbf{z}})$ the solution of the diffusion problem with data $\tilde{a} = a(\tilde{\mathbf{z}})$ and $\tilde{f} = f(\tilde{\mathbf{z}})$.
Subtract the weak form with the solution $\tilde{u}$ from the weak form with solution u. Then it is for all $v \in H^1_0(D)$

$$\begin{aligned}
\int_D (f - \tilde{f})v \,\mathrm{d}\mathbf{x} &= \int_D a\nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} - \int_D \tilde{a}\nabla\tilde{u} \cdot \nabla v \,\mathrm{d}\mathbf{x} \\
&= \int_D a(\nabla u - \nabla\tilde{u}) \cdot \nabla v \,\mathrm{d}\mathbf{x} + \int_D a\nabla\tilde{u} \cdot \nabla v \,\mathrm{d}\mathbf{x} - \int_D \tilde{a}\nabla\tilde{u} \cdot \nabla v \,\mathrm{d}\mathbf{x} \\
&= \int_D a\nabla(u - \tilde{u}) \cdot \nabla v \,\mathrm{d}\mathbf{x} + \int_D (a - \tilde{a})\nabla\tilde{u} \cdot \nabla v \,\mathrm{d}\mathbf{x}.
\end{aligned}$$

The function $w = u - \tilde{u}$ is the solution of

$$\int_D a\nabla w \cdot \nabla v \,\mathrm{d}\mathbf{x} = \int_D (f - \tilde{f})v \,\mathrm{d}\mathbf{x} + \int_D (\tilde{a} - a)\nabla\tilde{u} \cdot \nabla v \,\mathrm{d}\mathbf{x} = L(v)$$

and the stability estimate

$$\|w\|_{H^1_0(D)} \leq \frac{\|L\|_{H^1_0(D)^*}}{\bar{a}_{min}}$$

holds. Since

$$\|L\|_{H_0^1(D)^*} = \max_{\|v\|_{H_0^1(D)}=1} |L(v)| \leq \|a - \tilde{a}\|_{L^\infty(D)} \|\tilde{u}\|_{H_0^1(D)} + C_P \|f - \tilde{f}\|_{L^2(D)}$$

and by Theorem 3.21 the stability estimate (3.6)

$$\|\tilde{u}\|_{H_0^1(D)} \leq \frac{C_P}{\bar{a}_{min}} \|\tilde{f}\|_{L^2(D)}$$

holds, the following inequality is obtained:

$$\|L\|_{H_0^1(D)^*} \leq \|a - \tilde{a}\|_{L^\infty(D)} \frac{C_P \|\tilde{f}\|_{L^2(D)}}{\bar{a}_{min}} + C_P \|f - \tilde{f}\|_{L^2(D)}.$$

This gives

$$\|w\|_{H_0^1(D)} \leq \frac{\|L\|_{H_0^1(D)^*}}{\bar{a}_{min}} \leq \|a - \tilde{a}\|_{L^\infty(D)} \frac{C_P \|\tilde{f}\|_{L^2(D)}}{\bar{a}_{min}^2} + \frac{C_P \|f - \tilde{f}\|_{L^2(D)}}{\bar{a}_{min}},$$

and for all $\mathbf{z}, \tilde{\mathbf{z}} \in \Sigma_U$ such that $|\mathbf{z} - \tilde{\mathbf{z}}| \leq \delta$ it holds by (3.7) and (3.8)

$$\|u(\mathbf{z}) - u(\tilde{\mathbf{z}})\|_{H_0^1(D)} = \|w\|_{H_0^1(D)} \leq \epsilon,$$

and hence the solution u is continuous in Σ_U. With (3.6), $\sigma(\operatorname{Re}\mathbf{z})$ constant with respect to $\|\cdot\|_{H_0^1(D)}$, and $f \in C_\sigma^0(\Sigma_U; L^2(D))$ it follows

$$\begin{aligned}
\|u\|_{C_\sigma^0(\Sigma_U; H_0^1(D))} &= \max_{\mathbf{z} \in \Sigma_U} \|\sigma(\operatorname{Re}\mathbf{z}) u(\mathbf{z})\|_{H_0^1(D)} \leq \max_{\mathbf{z} \in \Sigma_U} \left(\sigma(\operatorname{Re}\mathbf{z}) \frac{C_P}{\bar{a}_{min}} \|f(\mathbf{z})\|_{L^2(D)} \right) \\
&\leq \frac{C_P}{\bar{a}_{min}} \|f\|_{C_\sigma^0(\Sigma_U; L^2(D))} < \infty.
\end{aligned}$$

Hence, $u \in C_\sigma^0(\Sigma_U; L^2(D))$. ■

3.3.1.2 Mixed form of the diffusion problem

Now, step 3 of Remark 3.11 is performed for the uniformly bounded diffusion coefficient. Begin with step $3a$.

Assumption 3.26 – Data in the mixed form:
For some given density function $\rho : \Sigma_U \to \mathbb{R}_+$ the data f and g should fulfil

- $\rho(\mathbf{z})\,\mathrm{d}\mathbf{z}$-a.e. it is $f(\mathbf{z}) \in L^2(D)$, where $\mathbf{z} \in \Sigma_U$, and
- $g = 0$. □

Remark 3.27:
Note that these assumptions are the same as for the diffusion problem, see Assumption 3.19. Only the assumption on the boundary condition is added. For the diffusion problem in standard form the boundary condition was contained in the assumption on the spaces (see Assumption 3.16). □

Introduce the following space (recall that $d = \dim(D)$)

$$H(\mathrm{div}; D) = \{\mathbf{v} \in [L^2(D)]^d : \nabla \cdot \mathbf{v} \in L^2(D)\}$$

equipped with the norm

$$\|\mathbf{v}\|_{H(\mathrm{div};D)} = \sqrt{\|\mathbf{v}\|^2_{L^2(D)} + \|\nabla \cdot \mathbf{v}\|^2_{L^2(D)}}.$$

Assumption 3.28 – Spaces of the mixed form:
Let $W(D) = L^2(D) \times H(\mathrm{div}; D)$ and again $q = 2$. Then it is

$$L^q_\rho(\Gamma; W(D)) = L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D)).$$

□

The weak formulation (extended to $\mathbb{C}^N$) of the diffusion problem in mixed form reads: Find a function $(p(\mathbf{z}), \mathbf{u}(\mathbf{z})) \in L^2(D) \times H(\mathrm{div}; D)$ such that for any $(q, \mathbf{v}) \in L^2(D) \times H(\mathrm{div}; D)$ $\rho(\mathbf{z})\,\mathrm{d}\mathbf{z}$-a.e. it holds

$$\begin{aligned} \int_D \frac{1}{a(\mathbf{z},\mathbf{x})} \mathbf{u}(\mathbf{z},\mathbf{x})\mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_D p(\mathbf{z},\mathbf{x}) \nabla \cdot \mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x} &= 0, \\ \int_D q(\mathbf{x}) \nabla \cdot \mathbf{u}(\mathbf{z},\mathbf{x})\,\mathrm{d}\mathbf{x} &= -\int_D f(\mathbf{z},\mathbf{x}) q(\mathbf{x})\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{3.9}$$

Define

$$\begin{aligned} A_{\mathbf{z}}(\mathbf{u}, \mathbf{v}) &= \int_D \frac{1}{a(\mathbf{z},\mathbf{x})} \mathbf{u}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x}, \\ B(\mathbf{v}, q) &= \int_D q(\mathbf{x}) \nabla \cdot \mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x}, \\ h_{\mathbf{z}}(q) &= -\int_D f(\mathbf{z},\mathbf{x}) q(\mathbf{x})\,\mathrm{d}\mathbf{x}. \end{aligned}$$

Equation (3.9) is equivalent to

$$\begin{aligned} A_{\mathbf{z}}(\mathbf{u}(\mathbf{z},\cdot), \mathbf{v}) + B(\mathbf{v}, p(\mathbf{z},\cdot)) &= 0 && \forall \mathbf{v} \in H(\mathrm{div}; D), \\ B(\mathbf{u}(\mathbf{z},\cdot), q) &= h_{\mathbf{z}}(q) && \forall q \in L^2(D). \end{aligned}$$

Theorem 3.29 – Existence of a unique solution of (3.9):
With Assumptions 3.10, 3.26, and 3.28 on the extension of the data, the data itself and on the spaces of the mixed form, for all $\mathbf{z} \in \Sigma_U$ *there exists a unique solution* $(p(\mathbf{z}), \mathbf{u}(\mathbf{z})) \in L^2(D) \times H(\mathrm{div}; D)$ *of (3.9) satisfying* $\rho(\mathbf{z})\, d\mathbf{z}$*-a.e.*

$$\|\mathbf{u}(\mathbf{z})\|_{H(\mathrm{div};D)} \leq C\left(\frac{\bar{a}_{max}}{\bar{a}_{min}}\right)\|f(\mathbf{z})\|_{L^2(D)},$$
$$\|p(\mathbf{z})\|_{L^2(D)} \leq C^2\left(\frac{\bar{a}_{max}}{\bar{a}_{min}^2}\right)\|f(\mathbf{z})\|_{L^2(D)}$$

with a constant $C > 0$ *depending on the domain only.* □

Proof: The proof is similar to [5, pp. 2045–2046] (where it is only stated for a deterministic force term and a general boundary condition g) and almost the same as in [11, pp. 6–7] (there for the D-bounded case). For the proof it will be shown that all conditions of Theorem B.10 (inf-sup condition for the existence and uniqueness of a saddle point problem) are fulfilled with $\mathbf{u},\ \mathbf{v} \in V = H(\mathrm{div}; D)$ and $p, q \in Q = L^2(D)$.

1. $A_{\mathbf{z}}$ and B are continuous linear forms and $h_{\mathbf{z}}$ is a bounded linear functional.
It is $\|\frac{1}{a(\mathbf{z})}\|_{L^\infty(D)} \leq \frac{1}{\bar{a}_{min}}$ because for all $\mathbf{z} \in \Sigma_U$ it holds $\bar{a}_{min} \leq a(\mathbf{z}, \mathbf{x})$ for all $\mathbf{x} \in D$. Hence,

$$\begin{aligned}
|A_{\mathbf{z}}(\mathbf{u}, \mathbf{v})| &= \left|\left(\tfrac{1}{a(\mathbf{z})}\mathbf{u}, \mathbf{v}\right)\right| \leq \left\|\frac{1}{a(\mathbf{z})}\right\|_{L^\infty(D)} \|\mathbf{u}\|_{H(\mathrm{div};D)}\|\mathbf{v}\|_{H(\mathrm{div};D)} \\
&\leq \frac{1}{\bar{a}_{min}}\|\mathbf{u}\|_{H(\mathrm{div};D)}\|\mathbf{v}\|_{H(\mathrm{div};D)}, \\
|B(\mathbf{v}, q)| &= |(q, \nabla \cdot \mathbf{v})| \leq \|q\|_{L^2(D)}\|\mathbf{v}\|_{H(\mathrm{div};D)}, \\
|h_{\mathbf{z}}(q)| &= |(f(\mathbf{z}), q)| \leq \|f(\mathbf{z})\|_{L^2(D)}\|q\|_{L^2(D)}.
\end{aligned}$$

Note that the continuity of $A_{\mathbf{z}}$ and B follows from the boundedness and the linearity from the linearity of the integrals. Further, it is $\|A_{\mathbf{z}}\| = \frac{1}{\bar{a}_{min}}$.

2. The inf-sup condition holds.
Define the space

$$\begin{aligned}
V^0 &= \{\mathbf{v} \in H(\mathrm{div}; D) : B(\mathbf{v}, q) = (q, \nabla \cdot \mathbf{v}) = 0 \ \ \forall q \in L^2(D)\} \\
&= \{\mathbf{v} \in H(\mathrm{div}; D) : \|\nabla \cdot \mathbf{v}\|_{L^2(D)} = 0\}.
\end{aligned}$$

Then (see [18, pp. 40–41] or [11, p. 6]), for each q in $L^2(D)$ there exists a unique $\mathbf{v}_q$ in $V^{0,\perp}$, the orthogonal complement of V^0, where the orthogonality is understood with respect to the inner product of $H(\mathrm{div}; D)$, such that

$$\nabla \cdot \mathbf{v}_q = q \quad \text{and} \quad \|\mathbf{v}_q\|_{H(\mathrm{div};D)} \leq C\|q\|_{L^2(D)}.$$

The constant C depends only on the domain D. Note that $V^{0,\perp} \subset H(\mathrm{div}; D)$ because of the orthogonal decomposition

$$V^0 \oplus V^{0,\perp} = H(\mathrm{div}; D)$$

of the space $H(\mathrm{div}; D)$, and hence $\mathbf{v}_q \in H(\mathrm{div}; D)$.
It follows

$$\begin{aligned}
\sup_{\mathbf{v}\in H(\mathrm{div};D)} \frac{B(\mathbf{v}, q)}{\|\mathbf{v}\|_{H(\mathrm{div};D)}} &= \sup_{\mathbf{v}\in H(\mathrm{div};D)} \frac{(q, \nabla \cdot \mathbf{v})}{\|\mathbf{v}\|_{H(\mathrm{div};D)}} \\
&\overset{\mathbf{v}=\mathbf{v}_q}{\geq} \frac{(q, \nabla \cdot \mathbf{v}_q)}{\|\mathbf{v}_q\|_{H(\mathrm{div};D)}} \overset{\nabla\cdot\mathbf{v}_q=q}{=} \frac{\|q\|^2_{L^2(D)}}{\|\mathbf{v}_q\|_{H(\mathrm{div};D)}} \geq \frac{1}{C}\|q\|_{L^2(D)}.
\end{aligned}$$

Note that the constant k_0 from Theorem B.10 is given by $k_0 = \frac{1}{C}$.

3. A is coercive on V^0.
For all $\mathbf{z} \in \Sigma_U$ it is $\frac{1}{\bar{a}_{max}} \leq \frac{1}{a(\mathbf{z},\mathbf{x})}$ because $a(\mathbf{z}, \mathbf{x}) \leq \bar{a}_{max}$. Hence, it is $\operatorname{ess\,inf}_{\mathbf{x}\in D} \frac{1}{a(\mathbf{z},\mathbf{x})} \geq \frac{1}{\bar{a}_{max}}$.
Let $\mathbf{v} \in V^0$. Then,

$$\begin{aligned}
A_{\mathbf{z}}(\mathbf{v}, \mathbf{v}) &= \left(\frac{1}{a(\mathbf{z})}\mathbf{v}, \mathbf{v}\right) \\
&\geq \operatorname*{ess\,inf}_{\mathbf{x}\in D} \frac{1}{a(\mathbf{z}, \mathbf{x})}\|\mathbf{v}\|^2_{L^2(D)} \geq \frac{1}{\bar{a}_{max}}\|\mathbf{v}\|^2_{L^2(D)} = \frac{1}{\bar{a}_{max}}\|\mathbf{v}\|^2_{H(\mathrm{div};D)},
\end{aligned}$$

where the last equality holds due to

$$\|\mathbf{v}\|^2_{H(\mathrm{div};D)} = \|\mathbf{v}\|^2_{L^2(D)} + \|\nabla \cdot \mathbf{v}\|^2_{L^2(D)} = \|\mathbf{v}\|^2_{L^2(D)} \quad \text{for } \mathbf{v} \in V^0.$$

For the constant α_0 from Theorem B.10 it holds $\alpha_0 = \frac{1}{\bar{a}_{max}}$.
Since all conditions of Theorem B.10 are fulfilled, a unique solution for each $\mathbf{z} \in \Sigma_U$ is obtained.
The stability estimates follow by inserting the calculated constants in the

estimates given in Theorem B.10 and using $\frac{\bar{a}_{min}}{\bar{a}_{max}} \leq 1$

$$\begin{aligned}
\|p\|_{L^2(D)} &\leq \frac{\|A\|}{k_0^2}\left(\frac{\|A\|}{\alpha_0}+1\right)\|f\|_{L^2(D)} \\
&= \frac{1}{\bar{a}_{min}}C^2\left(\frac{1}{\bar{a}_{min}}\bar{a}_{max}+1\right)\|f\|_{L^2(D)} \\
&= \frac{\bar{a}_{max}}{\bar{a}_{min}^2}C^2\left(1+\frac{\bar{a}_{min}}{\bar{a}_{max}}\right)\|f\|_{L^2(D)} \\
&\leq \frac{\bar{a}_{max}}{\bar{a}_{min}^2}\tilde{C}^2\|f\|_{L^2(D)}
\end{aligned}$$

and

$$\begin{aligned}
\|\mathbf{u}\|_{H(\mathrm{div};D)} &\leq \left(\frac{\|A\|}{\alpha_0}+1\right)\frac{1}{k_0}\|f\|_{L^2(D)} \\
&= \left(\frac{\bar{a}_{max}}{\bar{a}_{min}}+1\right)C\|f\|_{L^2(D)} \\
&= \frac{\bar{a}_{max}}{\bar{a}_{min}}C\left(1+\frac{\bar{a}_{min}}{\bar{a}_{max}}\right)\|f\|_{L^2(D)} \\
&\leq \frac{\bar{a}_{max}}{\bar{a}_{min}}\tilde{C}\|f\|_{L^2(D)}. \qquad \blacksquare
\end{aligned}$$

Due to step $3b$ of Remark 3.11, it has to be shown that the solution $(p, \mathbf{u})$ is contained in $C^0_\sigma(\Sigma_U; L^2(D) \times H(\mathrm{div}; D))$. Note that the following assumption is the same (except it is stated on $\frac{1}{a}$ and not on a) as the one for the diffusion problem.

Assumption 3.30 – Continuous data in the mixed form:
Let $f \in C^0_\sigma(\Sigma_U; L^2(D))$, $\frac{1}{a} \in C^0(\Sigma_U; L^\infty(D))$. $\square$

Lemma 3.31 – Continuous solution of (3.9):
With the assumptions of Theorem 3.29 and with Assumption 3.30 the solution $(p(\mathbf{z}), \mathbf{u}(\mathbf{z}))$ of the mixed form (3.9) is in $C^0_\sigma(\Sigma_U; L^2(D) \times H(\mathrm{div}; D))$, i.e., $(p, \mathbf{u}) \in C^0_\sigma(\Sigma_U; L^2(D) \times H(\mathrm{div}; D))$. $\square$

Proof: The continuity follows from the continuity of the data and can be shown by a similar argument as in [7, pp. 20–21], which has been given in the proof of Lemma 3.25. The result follows by estimating (using the stability estimates given in Theorem 3.29)

$$\begin{aligned}
\|p\|_{C^0_\sigma(\Sigma_U;L^2(D))} &= \max_{\mathbf{z}\in\Sigma_U}\|\sigma(\operatorname{Re}\mathbf{z})p(\mathbf{z})\|_{L^2(D)} = \max_{\mathbf{z}\in\Sigma_U}\sigma(\operatorname{Re}\mathbf{z})\|p(\mathbf{z})\|_{L^2(D)}\\
&\overset{3.29}{\leq} C^2\left(\frac{\bar{a}_{max}}{\bar{a}^2_{min}}\right)\max_{\mathbf{z}\in\Sigma_U}\sigma(\operatorname{Re}\mathbf{z})\|f(\mathbf{z})\|_{L^2(D)}\\
&\leq C^2\left(\frac{\bar{a}_{max}}{\bar{a}^2_{min}}\right)\|f\|_{C^0_\sigma(\Sigma_U;L^2(D))}\\
&< \infty,
\end{aligned}$$

which is finite by the assumptions on the data, and in the same way it is

$$\|\mathbf{u}\|_{C^0_\sigma(\Sigma_U;H(\mathrm{div};D))} < \infty. \qquad \blacksquare$$

3.3.2 D-bounded diffusion coefficient in $\mathbb{C}^N$

Throughout this subsection the following assumption should hold.

Assumption 3.32 – D-bounded diffusion coefficient:
Let the diffusion coefficient be D-bounded (see (3.4)), i.e., $\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}$-a.e. it holds

$$0 < a_{min}(\mathbf{y}) = \operatorname*{ess\,inf}_{\mathbf{x}\in D} a(\mathbf{y},\mathbf{x}) \leq \operatorname*{ess\,sup}_{\mathbf{x}\in D} a(\mathbf{y},\mathbf{x}) = a_{max}(\mathbf{y}) < \infty.$$

□

Once more, the line of reasoning is started with step 1 of Remark 3.11 by defining relevant subspaces of the complex plane.

Definition 3.33 – The set Σ_D:
Let the set Σ_D – the index $_D$ denotes the D-bounded case – be defined as

$$\Sigma_D = \{\mathbf{z}\in\mathbb{C}^N : 0 < \bar{a}_{min}(\mathbf{z}) \leq \operatorname{Re} a(\mathbf{z},\mathbf{x}) \leq \bar{a}_{max}(\mathbf{z}) < \infty\},$$

where

$$\bar{a}_{min}(\mathbf{z}) = \operatorname*{ess\,inf}_{\mathbf{x}\in D}\operatorname{Re} a(\mathbf{z},\mathbf{x}),\ \ \bar{a}_{max}(\mathbf{z}) = \operatorname*{ess\,sup}_{\mathbf{x}\in D}\operatorname{Re} a(\mathbf{z},\mathbf{x}).$$

□

The space $C^0_\sigma(\Gamma;W(D))$ is extended to functions defined from Σ_D to $W(D)$.

Definition 3.34 – The space $C^0_\sigma(\Sigma_D; W(D))$:
The space $C^0_\sigma(\Sigma_D; W(D)) =$

$$\left\{\mathbf{v} : \Sigma_D \to W(D) : \mathbf{v} \textit{ continuous in } \mathbf{z}, \max_{\mathbf{z}\in\Sigma_D} \|\sigma(\operatorname{Re}\mathbf{z})\mathbf{v}(\mathbf{z})\|_{W(D)} < \infty\right\}$$

$$\textit{with norm } \|\mathbf{v}\|_{C^0_\sigma(\Sigma_D;W(D))} = \max_{\mathbf{z}\in\Sigma_D} \|\sigma(\operatorname{Re}\mathbf{z})\mathbf{v}(\mathbf{z})\|_{W(D)}$$

consists of continuous functions $\mathbf{v} : \Sigma_D \to W(D)$ *whose norm* $\|\cdot\|_{C^0_\sigma(\Sigma_D;W(D))}$ *is finite.* □

Remark 3.35 – $\Gamma \subset \Sigma_D$:
$\Gamma \subset \Sigma_D$ holds in the case of a D-bounded diffusion coefficient as in Assumption 3.32. It is $\operatorname{Re} a(\mathbf{y}, \mathbf{x}) = a(\mathbf{y}, \mathbf{x})$ for all $\mathbf{y} \in \Gamma$, and $\bar{a}_{min}(\mathbf{y}) = a_{min}(\mathbf{y})$ by the equalities $\bar{a}_{min}(\mathbf{y}) = \operatorname{ess\,inf}_{\mathbf{x}\in D} \operatorname{Re} a(\mathbf{y}, \mathbf{x}) = \operatorname{ess\,inf}_{\mathbf{x}\in D} a(\mathbf{y}, \mathbf{x}) = a_{min}(\mathbf{y})$. Similarly it holds $\bar{a}_{max}(\mathbf{y}) = a_{max}(\mathbf{y})$. Thus, each $\mathbf{y} \in \Gamma$ is an element of Σ_D. □

3.3.2.1 Diffusion problem

The same procedure employed for the diffusion problem with a uniformly bounded diffusion coefficient will be repeated, i.e., the two parts of step 2 of Remark 3.11 will be carried out. Begin with step 2a.

Theorem 3.36 – Existence of a unique solution of (3.5):
Let Assumptions 3.10 and 3.32 on the extension of the data to $\mathbb{C}^N$ *and on the diffusion coefficient and Assumption 3.16 on the spaces and Assumption 3.19 on the force term hold. Then, for every fixed* $\mathbf{z} \in \Sigma_D$ *the diffusion problem in weak form (3.5) admits a unique solution* $u(\mathbf{z})$ *in* $H^1_0(D)$ *such that* $\rho(\mathbf{z})\,\mathrm{d}\mathbf{z}$ *- a.e.*

$$\|u(\mathbf{z})\|_{H^1_0(D)} \leq \frac{C_P}{\bar{a}_{min}(\mathbf{z})} \|f(\mathbf{z})\|_{L^2(D)} \tag{3.10}$$

with the Poincaré constant C_P *depending only on the domain* D. □

Proof: A complex valued version of the Lemma of Lax-Milgram B.8 has to be applied again. The stability estimate follows by the Poincaré inequality B.9. ■

For step 2b of Remark 3.11 the following assumption is stated.

Assumption 3.37 – Continuous data:
Let $f \in C^0_{\sqrt{\sigma}}(\Sigma_D; L^2(D))$, $\frac{1}{\bar{a}_{min}} \in C^0_{\sqrt{\sigma}}(\Sigma_D; \mathbb{R})$, and $a \in C^0(\Sigma_D; L^\infty(D))$. □

Note that this assumption is stricter than Assumption 3.24 on the force term in Σ_U. Since $\sigma \leq \sqrt{\sigma}$ by the assumption $\sigma \leq 1$ on the weight function, and if $f \in C^0_{\sqrt{\sigma}}(\Sigma_D; L^2(D))$, it follows $f \in C^0_\sigma(\Sigma_D; L^2(D))$ from

$$\begin{aligned}\|f\|_{C^0_\sigma(\Sigma_D;L^2(D))} &= \max_{\mathbf{z}\in\Sigma_D} \sigma(\operatorname{Re}\mathbf{z})\|f\|_{L^2(D)} \\ &\leq \max_{\mathbf{z}\in\Sigma_D} \sqrt{\sigma(\operatorname{Re}\mathbf{z})}\|f\|_{L^2(D)} = \|f\|_{C^0_{\sqrt{\sigma}}(\Sigma_D;L^2(D))}.\end{aligned}$$

Additionally, an assumption on $\frac{1}{\bar{a}_{min}}$ is stated. These differences are based on the different forms of the diffusion coefficient.

Lemma 3.38 – Continuous solution of (3.5):
Let the assumptions of the previous Theorem 3.36 and Assumption 3.37 hold. Then, it is $u \in C^0_\sigma(\Sigma_D; H^1_0(D))$. □

Proof: Similarly to the proof in the case of a uniformly bounded coefficient (see the proof of Lemma 3.25), $u \in C^0_\sigma(\Sigma_D; H^1_0(D))$ will be shown. The solution $u(\mathbf{z})$ is continuous with respect to $\mathbf{z} \in \Sigma_D$ by the same steps as in the proof of Lemma 3.25 using the stability estimate

$$\|u(\mathbf{z})\|_{H^1_0(D)} \leq \frac{C_P}{\bar{a}_{min}(\mathbf{z})}\|f(\mathbf{z})\|_{L^2(D)}$$

and the assumption on the data to be continuous with respect to $\mathbf{z} \in \Sigma_D$. Additionally, Assumption 3.37 on the data yields

$$\begin{aligned}\|u\|_{C^0_\sigma(\Sigma_D;H^1_0(D))} &= \max_{\mathbf{z}\in\Sigma_D}\|\sigma(\operatorname{Re}\mathbf{z})u(\mathbf{z})\|_{H^1_0(D)} = \max_{\mathbf{z}\in\Sigma_D}\sigma(\operatorname{Re}\mathbf{z})\|u(\mathbf{z})\|_{H^1_0(D)} \\ &\overset{(3.10)}{\leq} C_P \max_{\mathbf{z}\in\Sigma_D} \frac{\sqrt{\sigma(\operatorname{Re}\mathbf{z})}}{\bar{a}_{min}(\mathbf{z})}\left(\sqrt{\sigma(\operatorname{Re}\mathbf{z})}\|f(\mathbf{z})\|_{L^2(D)}\right) \\ &\leq C_P \left\|\frac{1}{\bar{a}_{min}}\right\|_{C^0_{\sqrt{\sigma}}(\Sigma_D;\mathbb{R})}\|f\|_{C^0_{\sqrt{\sigma}}(\Sigma_D;L^2(D))} < \infty.\end{aligned}$$ ■

3.3.2.2 Mixed form of the diffusion problem

The distinctions of cases are completed by the steps $3a$ and $3b$ for the mixed form.

Theorem 3.39 – Existence of a unique solution of (3.9):
Let Assumptions 3.10, 3.32, 3.26, and 3.28, i.e., the assumptions on the extendibility of the data to $\mathbb{C}^N$, the diffusion coefficient, the data itself and

the spaces of the mixed form, hold. Then, for each $\mathbf{z} \in \Sigma_D$ *there exists a unique solution* $(p(\mathbf{z}), \mathbf{u}(\mathbf{z}))$ *of (3.9) satisfying* $\rho(\mathbf{z})\,d\mathbf{z}$*-a.e.*

$$\|\mathbf{u}(\mathbf{z})\|_{H(\mathrm{div};D)} \leq C\left(\frac{\bar{a}_{max}(\mathbf{z})}{\bar{a}_{min}(\mathbf{z})}\right)\|f(\mathbf{z})\|_{L^2(D)},$$
$$\|p(\mathbf{z})\|_{L^2(D)} \leq C^2\left(\frac{\bar{a}_{max}(\mathbf{z})}{\bar{a}^2_{min}(\mathbf{z})}\right)\|f(\mathbf{z})\|_{L^2(D)}$$

with a constant $C > 0$ *depending on the domain only.* □

Proof: A proof can be found in [11, pp. 6–7]. It is almost identical to the procedure used for the uniformly bounded coefficient of Theorem 3.29, where $\bar{a}_{min}$ is replaced by $\bar{a}_{min}(\mathbf{z})$. ■

Step $3b$ needs the following continuity assumptions on the data.

Assumption 3.40 – Continuous data:
Assume that $f \in C^0_{\sqrt[4]{\sigma}}(\Sigma_D; L^2(D))$, $a \in C^0_{\sqrt[4]{\sigma}}(\Sigma_D; L^\infty(D))$, $\bar{a}_{max} \in C^0_{\sqrt[4]{\sigma}}(\Sigma_D; \mathbb{R})$, and $\frac{1}{\bar{a}_{min}} \in C^0_{\sqrt[4]{\sigma}}(\Sigma_D; \mathbb{R})$. □

Lemma 3.41 – Continuous solution of (3.9):
Let the assumptions of the previous Theorem 3.39 hold. Further, let Assumption 3.40 hold. Then, $(p, \mathbf{u}) \in C^0_\sigma(\Sigma_D; L^2(D) \times H(\mathrm{div}; D))$*.* □

Proof: A proof can be found in [11, p. 7]. Note that Assumption 3.40 is formulated here as to fulfil $(p, \mathbf{u}) \in C^0_\sigma(\Sigma_D; L^2(D) \times H(\mathrm{div}; D))$ while the assumptions given in [11, p. 6] are for σ^4 instead of σ. Again, continuity follows by the assumptions on the data. The result can be obtained by estimates similar to those in the previous proofs on continuity. ■

3.3.3 Results in $\mathbb{R}^N$

Not only the results in $\mathbb{C}^N$ are of interest, but also results in $\mathbb{R}^N$ ensuring existence and uniqueness of the diffusion problem and its mixed form in the space $L^2_\rho(\Gamma; W(D))$.
That is, a unique weak solution of both problems now stated in the form of (2.4) is sought. Both problem formulations are given for the sake of completeness.
Find a function $u \in L^2_\rho(\Gamma; H^1_0(D))$ such that for any $\mathbf{v} \in L^2_\rho(\Gamma; H^1_0(D))$ it holds

$$\int_\Gamma \int_D a(\mathbf{y}, \mathbf{x})\nabla u(\mathbf{y}, \mathbf{x})\nabla \mathbf{v}(\mathbf{y}, \mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \int_\Gamma \int_D f(\mathbf{y}, \mathbf{x})\mathbf{v}(\mathbf{y}, \mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \tag{3.11}$$

where $u(\mathbf{y},\mathbf{x}) = 0$ on $\Gamma \times \partial D$ has been used.
And: Find a function $(p,\mathbf{u}) \in L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$ such that for any $(q,\mathbf{v}) \in L^2_\rho(\Gamma; L^2(D)) \otimes L^2_\rho(\Gamma; H(\mathrm{div}; D))$ it holds

$$\begin{aligned}
&\int_\Gamma \int_D \frac{1}{a(\mathbf{y},\mathbf{x})}\mathbf{u}(\mathbf{y},\mathbf{x})\mathbf{v}(\mathbf{y},\mathbf{x}) - \nabla p(\mathbf{y},\mathbf{x})\mathbf{v}(\mathbf{y},\mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = 0\\
&\int_\Gamma \int_D \nabla \cdot \mathbf{u}(\mathbf{y},\mathbf{x})q(\mathbf{y},\mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \int_\Gamma \int_D -f(\mathbf{y},\mathbf{x})q(\mathbf{y},\mathbf{x})\rho(\mathbf{y})\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}
\end{aligned} \tag{3.12}$$

where $p(\mathbf{y},\mathbf{x}) = 0$ on $\Gamma \times \partial D$ has been used.

The results given in the previous two subsections for subsets of $\mathbb{C}^N$ defined by a uniformly bounded and a D-bounded diffusion coefficient can be understood in the subsets $\Gamma \subset \Sigma_U$ or $\Gamma \subset \Sigma_D$ (with the stricter bounds a_{min} and a_{max} as well as $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$). The results of the diffusion problem and its mixed form with uniformly bounded diffusion coefficient in $\Gamma \subset \mathbb{R}^N$ can be derived as in Section 3.3.1 and with D-bounded diffusion coefficient as in Section 3.3.2.
Hence, for each case (i.e., diffusion problem and its mixed form with uniformly bounded diffusion coefficient or with D-bounded diffusion coefficient) only the additional assumptions to obtain the respective solution in the space $L^2_\rho(\Gamma; W(D))$ have to be stated.

Assumption 3.42 – Integrability of the force term:
Let the force term satisfy that $f(\mathbf{y},\cdot)$ is square integrable with respect to $\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}$, i.e.,

$$\int_D \int_\Gamma \big(f(\mathbf{y},\mathbf{x})\big)^2 \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} < \infty.$$

This means $f \in L^2_\rho(\Gamma; L^2(D))$. □

Lemma 3.43 – Uniformly bounded $\boldsymbol{a}$, $\boldsymbol{u \in L^2_\rho(\Gamma; H^1_0(D))}$:
Let the result of Theorem 3.21 for $\Gamma \subset \Sigma_U$ hold. Let Assumption 3.42 on the force term hold. Then, the diffusion problem (3.11) has a unique weak solution $u \in L^2_\rho(\Gamma; H^1_0(D))$. □

Proof: The property $u \in L^2_\rho(\Gamma; H^1_0(D))$ follows by

$$\begin{aligned}
\|u\|^2_{L^2_\rho(\Gamma;H^1_0(D))} &= \int_\Gamma \|u(\mathbf{y})\|^2_{H^1_0(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \overset{(3.6)}{\leq} \int_\Gamma \frac{C_P^2}{a_{min}^2}\|f(\mathbf{y})\|^2_{L^2(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\\
&= \frac{C_P^2}{a_{min}^2}\int_\Gamma \|f(\mathbf{y})\|^2_{L^2(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} < \infty
\end{aligned}$$

because by assumption f is in $L^2_\rho(\Gamma; L^2(D))$. ■

Lemma 3.44 – Unif. bdd. a, $(p, \mathrm{u}) \in L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$:
Let the result of Theorem 3.29 for $\Gamma \subset \Sigma_U$ hold. Let Assumption 3.42 on the force term hold. Then, the unique weak solution $(p, \mathbf{u})$ of the diffusion problem's mixed form (3.12) is in $L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$. □

Proof: As earlier stated in Lemma 3.43 for the diffusion problem in standard form, by integration of $\|p\|^2_{L^2(D)}$ and $\|\mathbf{u}\|^2_{H(\mathrm{div};D)}$ over Γ it can be shown that the solution $(p, \mathbf{u})$ is in $L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$. ■

Assumption 3.45 – D-bounded a, data of standard form:
Let for the data hold $f \in L^4_\rho(\Gamma; L^2(D))$ and $\frac{1}{a_{min}} \in L^4_\rho(\Gamma; \mathbb{R})$. □

The former assumption ensures that the solution u is in the right space.

Lemma 3.46 – D-bounded a, $u \in L^2_\rho(\Gamma; H^1_0(D))$:
Let Assumption 3.45 and the assumptions on Theorem 3.36 hold for $\Gamma \subset \Sigma_D$. Then, the diffusion problem in weak form (3.11) has a unique weak solution $u \in L^2_\rho(\Gamma; H^1_0(D))$. □

Proof: By equation (3.10), the Cauchy-Schwarz inequality (see Lemma B.2) and Assumption 3.45 it is

$$\begin{aligned}
\int_\Gamma \|u(\mathbf{y})\|^2_{H^1_0(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} &\overset{(3.10)}{\leq} C_P \int_\Gamma \frac{1}{a_{min}(\mathbf{y})} \|f(\mathbf{y})\|^2_{L^2(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{\text{Lemma B.2}}{\leq} C_P \Big(\int_\Gamma \frac{\rho(\mathbf{y})}{a^4_{min}(\mathbf{y})}\,\mathrm{d}\mathbf{y} \Big)^{\frac{1}{2}} \Big(\int_\Gamma \|f(\mathbf{y})\|^4_{L^2(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \Big)^{\frac{1}{2}} \\
&\overset{\text{Ass. 3.45}}{<} \infty.
\end{aligned}$$

Therefore, $u \in L^2_\rho(\Gamma; H^1_0(D))$. ■

Assumption 3.47 – D-bounded a, data of mixed form:
Assume that $a \in L^s_\rho(\Gamma; L^\infty(D))$ for $s \in [1, \infty)$, $\frac{1}{a_{min}} \in L^s_\rho(\Gamma; \mathbb{R})$ for $s \in [1, \infty)$ and $f \in L^{q^*}_\rho(\Gamma; L^2(D))$ for some $q^* > 2$. □

Lemma 3.48 – D-bounded a, $(p, \mathrm{u}) \in L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$:
With Assumption 3.47 on the data and the assumptions of Theorem 3.39 for $\Gamma \subset \Sigma_D$ there exists a unique weak solution $(p, \mathbf{u}) \in L^2_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$ of (3.12). □

Proof: Since the statement is required on $\Gamma \subset \mathbb{R}^N$, it can be referred to the proof in [11, pp. 6–7] for $\Gamma \subset \mathbb{C}^N$ which is more general than $\Gamma \subset \mathbb{R}^N$. A proof with slightly different norms and estimates is given in [10, pp. 5–7]. ■

3.4 Analytic extension

As mentioned above, the analytic extension of the solution is required to bound the best approximation error occurring in the estimate of the stochastic approximation error (see Section 4.4) or to construct a quasi-optimal sparse grid. Subsequently, the analytic extension of the solution in one direction and in all directions simultaneously are given for the diffusion problem and its mixed form.
Denote

$$b(\mathbf{y}) = a(\mathbf{y}),\ \mathbf{y} \in \mathbb{R}^N \text{ and } b(\mathbf{z}) = a(\mathbf{z}),\ \mathbf{z} \in \mathbb{C}^N \tag{3.13}$$

in the diffusion problem and

$$b(\mathbf{y}) = \frac{1}{a(\mathbf{y})},\ \mathbf{y} \in \mathbb{R}^N \text{ and } b(\mathbf{z}) = \frac{1}{a(\mathbf{z})},\ \mathbf{z} \in \mathbb{C}^N \tag{3.14}$$

in the mixed form.
Further, some relevant notions of an analytic function are given.

Definition 3.49 – Analytic function of several variables:
A function $h : W \to \mathbb{C}$ is called analytic or holomorphic on an open domain $W \subset \mathbb{C}^N$ if there is for any point $w \in W$ an open region $U = U(w) \subset W$ and a power series which converges for all z in U to $h(z)$, i.e.,

$$h(z) = \sum_{\nu \geq 0} w_\nu (z - w)^\nu. \qquad \square$$

Remark 3.50:
For $N = 1$ the definition of an analytic function in only one variable is given. $\square$

Definition 3.51 – Partial differentiable and weakly holomorphic:
A function h is partial differentiable in w if the partial derivatives $\partial_n h(w)$ exist for $n \in \{1, \ldots, N\}$.
The function h is called weakly holomorphic in U if h is continuous in U and partial differentiable.
For $z = (z_1, \ldots, z_n) \in U$ and $n \in \{1, \ldots, N\}$ the mapping

$$\zeta \to h(z_1, \ldots, z_{n-1}, \zeta, z_{n+1}, \ldots, z_N)$$

is a holomorphic function of one variable. $\square$

Note that by the Theorem of Osgood (given in Theorem C.2) the notions of analytic/holomorphic and weakly holomorphic are equivalent.

Definition 3.52 – Analytic extension:
Let $W \subset \mathbb{C}^N$ be an open subset. A function $h : W \to \mathbb{C}$ has an analytic extension (or continuation) to the open domain $M \subset \mathbb{C}^N$ with $W \subset M$ if $\tilde{h} : M \to \mathbb{C}$ is analytic and $\tilde{h}(z) = h(z)\ \forall z \in W$. $\square$

3.4.1 Analytic extension – one-dimensional result

To obtain the one-dimensional analyticity result, the following notation is introduced:

$$\Gamma_n^* = \prod_{j=1, j\neq n}^{N} \Gamma_j, \qquad \sigma_n^* = \prod_{j=1, j\neq n}^{N} \sigma_j.$$

Assumption 3.53 – Differentiability of b and f:
Let $b(\mathbf{y})$ and the force term $f(\mathbf{y})$ be infinitely many times differentiable with respect to $\mathbf{y}$, i.e., for each direction y_n the k-th derivatives $\partial_{y_n}^k b(\mathbf{y})$ and $\partial_{y_n}^k f(\mathbf{y})$ exist for every $k \in \mathbb{N}_0$. Additionaly, if $b(\mathbf{y}) = \frac{1}{a(\mathbf{y})}$, let the input data $\frac{1}{a(\mathbf{y})}$ and $f(\mathbf{y})$ be analytic on Γ. □

For the diffusion problem and its mixed form the following statement holds (see [1, p. 1016] and [10, p. 7]).

Lemma 3.54 – Analytic extension of w in one dimension to $\mathbb{C}$:
Let Assumption 3.53 and the assumptions for existence and uniqueness of a continuous solution of the diffusion problem and its mixed form in $\Gamma \subset \mathbb{R}^N$ hold, i.e., the assumptions of Lemma 3.25 and 3.43, 3.31 and 3.44, 3.38 and 3.46, 3.41 and 3.48, respectively. Then, if for every $\mathbf{y} \in \Gamma$ there exists $\gamma_n < \infty$ such that for every $k \in \mathbb{N}_0$

$$\left\| \frac{\partial_{y_n}^k b(\mathbf{y})}{b(\mathbf{y})} \right\|_{L^\infty(D)} \leq \gamma_n^k k! \quad and \quad \frac{\|\partial_{y_n}^k f(\mathbf{y})\|_{L^2(D)}}{1 + \|f(\mathbf{y})\|_{L^2(D)}} \leq \gamma_n^k k!,$$

the solution $\mathbf{w}(\mathbf{y}, \mathbf{x}) = \mathbf{w}(y_1, \ldots, y_n, \ldots, y_N, \mathbf{x})$ as a function of y_n, i.e.,

$$\mathbf{w}|_n : \Gamma_n \to C^0_{\sigma_n^*}(\Gamma_n^*; W(D))$$

admits an analytic extension $\mathbf{w}(y_1, \ldots, y_{n-1}, z_n, y_{n+1}, \ldots y_N, \mathbf{x}) = \mathbf{w}|_n(z_n)$, $z_n \in \mathbb{C}$, in the region of the complex plane

$$\Sigma(\Gamma_n; \tau_n) = \{z_n \in \mathbb{C} : \operatorname{dist}(z_n, \Gamma_n) \leq \tau_n\}$$

with $0 < \tau_n < \frac{1}{2\gamma_n}$.
Moreover, if σ_n is chosen such that a growth condition of the form

$$\sigma_n(\operatorname{Re} z_n) \leq C_n(\tau_n)\sigma_n(y_n) \quad \forall\ |z_n - y_n| \leq \tau_n$$

holds for all $y_n \in \Gamma_n$, where the constant $C_n(\tau_n)$ depends on n and τ_n only, then the following bound on the solution

$$\|\sigma_n(\operatorname{Re} z_n)\mathbf{w}|_n(z_n)\|_{C^0_{\sigma_n^*}(\Gamma_n^*; W(D))} \leq C_P(\tau_n, a, f)$$

is satisfied, where the constant $C_P(\tau_n, a, f)$ depends on the domain D, the direction n, τ_n and the data a and f. □

Proof: A proof for the diffusion problem with uniformly bounded diffusion coefficient can be found in [1, pp. 1016 – 1017]. If the proof is slightly changed by considering $a_{min}(\mathbf{y})$ instead of a_{min} and uses the assumptions on the continuity of the data belonging to a D-bounded diffusion coefficient, the case of a D-bounded diffusion coefficient can be proved using analogous steps.
The steps of a proof for the mixed form with D-bounded diffusion coefficient can be found in [10, pp. 8 – 11]. The norms and therefore the stability estimates of the solution are slightly different there. If the diffusion coefficient is uniformly bounded, the proof can be slightly modified by taking the constants a_{min} and a_{max} instead of $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$ as well as the assumptions on the data belonging to this case. ■

Remark 3.55 – Analyticity result for Smolyak approximations:
As discussed in [11, p. 14] the previous result – allowing only analytic extensions to a subset of the complex plane with respect to one variable and not to several variables at the same time – is not enough to bound the errors when using a Smolyak sparse grid (see Section 4.1.3).
Therefore, the results of this section have to be "improved", as it has been done in [11] for the mixed formulation. The results related to the diffusion problem in its standard form are given following closely the procedure in [11]. □

3.4.2 Analytic extension – product subdomain result

3.4.2.1 Uniformly bounded diffusion coefficient

In order to derive the analyticity of the solution, analyticity assumptions on the data are necessary.

Assumption 3.56 – Analytic data – uniformly bdd. coefficient:
Let $b : \mathbb{C}^N \to L^\infty(D)$ be analytic on $\mathbb{C}^N$, and let $f : \mathbb{C}^N \to L^2(D)$ be analytic on $\mathrm{int}(\Sigma_U)$, the interior of Σ_U. □

Theorem 3.57 – Analytic extension of w to subdomain of $\mathbb{C}^N$:
Let Assumption 3.56 hold, assume that $\Gamma \subset \mathrm{int}(\Sigma_U)$*, and let the assumptions of Lemma 3.25 and 3.31, respectively hold, i.e., there exists a continuous solution* $\mathbf{w} \in C^0_\sigma(\Sigma_U; W(D))$ *of the diffusion problem and its mixed form. Further, let there exist a unique solution* $\mathbf{w} \in L^2_\rho(\Gamma; W(D))$*, i.e., let the assumptions of Lemma 3.43 and 3.44, respectively, hold.*
Then, the function $\mathbf{z} \mapsto \mathbf{w}(\mathbf{z})$ *for the diffusion problem and its mixed form, is analytic in* $\mathrm{int}(\Sigma_U)$*, hence* $\mathbf{w} : \Gamma \to W(D)$ *has an analytic extension to the space* $\mathrm{int}(\Sigma_U)$*.* □

Proof: The proof will be given for the diffusion problem. It follows closely the lines of [8, pp. 9–12], which is extended here for the case of a stochastic force term, and [11, pp. 7–8], where it is given for the mixed form of the diffusion problem with D-bounded coefficient. The two main steps are

1. $\mathrm{int}(\Sigma_U)$ is an open set.

2. $\mathbf{z} \mapsto u(\mathbf{z})$ is analytic in $\mathrm{int}(\Sigma_U)$.

The first step follows automatically because the interior of some set, here Σ_U, is defined as the biggest open set contained in Σ_U.

The second step, i.e., the analyticity of the map $\mathbf{z} \mapsto u(\mathbf{z})$, is subdivided into four steps (denoted by 2a - 2d). The goal is to show the existence of the partial derivatives $\partial_n u(\mathbf{z})$ for $n \in \{1, \ldots, N\}$. The solution $u(\mathbf{z})$ is continuous in Σ_U (see Lemma 3.25) and thus in $\mathrm{int}(\Sigma_U)$. By the Theorem of Osgood (see Theorem C.2) the analyticity of u as a function of all N complex variables follows.

Step 2a: Define difference quotient.
Let $\mathbf{z} \in \mathrm{int}(\Sigma_U)$, $h \in \mathbb{C} \setminus \{0\}$ and $\mathbf{e}_n$ be the n-th unity vector in $\mathbb{R}^N$. Consider the difference quotient

$$q_h(\mathbf{z}) = \frac{u(\mathbf{z} + h\mathbf{e}_n) - u(\mathbf{z})}{h}. \tag{3.15}$$

Since the set $\mathrm{int}(\Sigma_U)$ is an open set, for every $\mathbf{z} \in \mathrm{int}(\Sigma_U)$ there is $\epsilon_{\mathbf{z}} > 0$ such that $\mathbf{z} + h\mathbf{e}_n \in \mathrm{int}(\Sigma_U)$ for $|h| < \epsilon_{\mathbf{z}}$. Thus, $u(\mathbf{z} + h\mathbf{e}_n)$ exists and the quotient is well defined if h is chosen small enough.

Step 2b: Define linear equation.
Denote by $A_{\mathbf{z}} : H_0^1(D) \to H_0^1(D)^*$ the linear mapping defined by

$$[A_{\mathbf{z}} u](v) = (a(\mathbf{z}) \nabla u(\mathbf{z}), \nabla v) \tag{3.16}$$

and by $f_{\mathbf{z}}$ a linear functional with $f_{\mathbf{z}}(v) = (f(\mathbf{z}), v)$. These definitions give

$$A_{\mathbf{z}} u = f_{\mathbf{z}}. \tag{3.17}$$

Let $\mathbf{z}_h = \mathbf{z} + h\mathbf{e}_n$. Then,

$$\begin{aligned}
[A_{\mathbf{z}} q_h](v) &\overset{(3.15)}{=} \left(a(\mathbf{z}) \nabla \frac{u(\mathbf{z} + h\mathbf{e}_n)}{h}, \nabla v \right) - \left(a(\mathbf{z}) \nabla \frac{u(\mathbf{z})}{h}, \nabla v \right) \\
&\overset{(3.16)}{=} \left(a(\mathbf{z}) \nabla \frac{u(\mathbf{z}_h)}{h}, \nabla v \right) - \frac{1}{h}[A_{\mathbf{z}} u](v) \\
&\overset{(3.17)}{=} \frac{1}{h}(a(\mathbf{z}) \nabla u(\mathbf{z}_h), \nabla v) - \frac{1}{h} f_{\mathbf{z}}(v) \\
&= \frac{1}{h}(a(\mathbf{z}) \nabla u(\mathbf{z}_h), \nabla v) - \frac{1}{h} f_{\mathbf{z}}(v) \\
&\quad \underbrace{- \frac{1}{h}(a(\mathbf{z}_h) \nabla u(\mathbf{z}_h), \nabla v) + \frac{1}{h}(a(\mathbf{z}_h) \nabla u(\mathbf{z}_h), \nabla v)}_{=0} \\
&= \frac{1}{h}(-[a(\mathbf{z}_h) - a(\mathbf{z})] \nabla u(\mathbf{z}_h), \nabla v) + \frac{1}{h}\big((a(\mathbf{z}_h) \nabla u(\mathbf{z}_h), \nabla v) - f_{\mathbf{z}}(v)\big) \\
&\overset{(3.17)}{=} \frac{1}{h}(-[a(\mathbf{z}_h) - a(\mathbf{z})] \nabla u(\mathbf{z}_h), \nabla v) + \frac{1}{h}(f_{\mathbf{z}_h}(v) - f_{\mathbf{z}}(v)) \\
&\overset{(3.16)}{=} \frac{-[(A_{\mathbf{z}_h} - A_{\mathbf{z}}) u(\mathbf{z}_h)](v)}{h} + \frac{f_{\mathbf{z}_h}(v) - f_{\mathbf{z}}(v)}{h}.
\end{aligned}$$

Let

$$L_h = \frac{-(A_{\mathbf{z}_h} - A_{\mathbf{z}}) u(\mathbf{z}_h)}{h} + \frac{f_{\mathbf{z}_h} - f_{\mathbf{z}}}{h}.$$

Consequently, q_h solves equation (3.17) for the right-hand side L_h.

Step 2c: Show that $L_h \to L_0$ as $h \to 0$.
Let

$$L_0 = -\partial_n A_{\mathbf{z}} u(\mathbf{z}) + \partial_n f_{\mathbf{z}},$$

where $[\partial_n A_{\mathbf{z}} \nabla u](v) = (\partial_n a(\mathbf{z}) \nabla u, \nabla v)$ and $\partial_n f_{\mathbf{z}}(v) = (\partial_n f(\mathbf{z}), v)$.
It will be shown that

$$\begin{aligned}
\lim_{h \to 0} \|L_h - L_0\|_{H_0^1(D)^*} \leq \lim_{h \to 0} &\left\| \frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} u(\mathbf{z}_h) - \partial_n A_{\mathbf{z}} u(\mathbf{z}) \right\|_{H_0^1(D)^*} \\
&+ \lim_{h \to 0} \left\| \frac{f_{\mathbf{z}_h} - f_{\mathbf{z}}}{h} - \partial_n f_{\mathbf{z}} \right\|_{H_0^1(D)^*} = 0.
\end{aligned}$$

Since f is analytic in $\mathrm{int}(\Sigma_U)$, it is

$$\partial_n f_{\mathbf{z}} = \lim_{h \to 0} \frac{f_{\mathbf{z}_h} - f_{\mathbf{z}}}{h} \in L^2(D)^* \subset H_0^1(D)^*,$$

and it follows with Remark 3.23 (as $L^2(D)^* = L^2(D)$ and $H_0^1(D)^* = H^{-1}(D)$ – see Remark B.7)

$$\lim_{h\to 0}\left\|\frac{f_{\mathbf{z}_h} - f_{\mathbf{z}}}{h} - \partial_n f_{\mathbf{z}}\right\|_{H_0^1(D)^*} \le C_P \lim_{h\to 0}\left\|\frac{f_{\mathbf{z}_h} - f_{\mathbf{z}}}{h} - \partial_n f_{\mathbf{z}}\right\|_{L^2(D)^*} = 0.$$

Further, it is

$$\begin{aligned}
&\left\|\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} u(\mathbf{z}_h) - \partial_n A_{\mathbf{z}} u(\mathbf{z})\right\|_{H_0^1(D)^*} \\
&= \left\|\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} u(\mathbf{z}_h) + \frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h}(-u(\mathbf{z}) + u(\mathbf{z})) - \partial_n A_{\mathbf{z}} u(\mathbf{z})\right\|_{H_0^1(D)^*} \\
&\le \left\|\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h}(u(\mathbf{z}_h) - u(\mathbf{z}))\right\|_{H_0^1(D)^*} + \left\|\left(\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} - \partial_n A_{\mathbf{z}}\right) u(\mathbf{z})\right\|_{H_0^1(D)^*} \\
&\le \left\|\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h}\right\| \|u(\mathbf{z}_h) - u(\mathbf{z})\|_{H_0^1(D)} + \left\|\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} - \partial_n A_{\mathbf{z}}\right\| \|u(\mathbf{z})\|_{H_0^1(D)}.
\end{aligned} \tag{3.18}$$

The first term in (3.18) tends to 0: Since $u \in C_\sigma^0(\Sigma_U; H_0^1(D))$, i.e., the solution is by assumption continuous in $\mathbf{z} \in \Sigma_U$, thus also in $\mathbf{z} \in \operatorname{int}(\Sigma_U)$, it is

$$\lim_{h\to 0}\|u(\mathbf{z}_h) - u(\mathbf{z})\|_{H_0^1(D)} = 0.$$

Additionally,

$$\begin{aligned}
\left\|\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h}\right\| &= \sup_{u,v\in H_0^1(D), u,v\neq 0} \frac{\left\|[\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} u](v)\right\|_{H_0^1(D)}}{\|u\|_{H_0^1(D)}\|v\|_{H_0^1(D)}} \\
&\le \sup_{u,v\in H_0^1(D), u,v\neq 0} \frac{\int_D \left|\frac{a(\mathbf{z}_h) - a(\mathbf{z})}{h} \nabla u \nabla v\right| \mathrm{d}\mathbf{x}}{\|u\|_{H_0^1(D)}\|v\|_{H_0^1(D)}} \\
&\le \left\|\frac{a(\mathbf{z}_h) - a(\mathbf{z})}{h}\right\|_{L^\infty(D)} \to \|\partial_n a(\mathbf{z})\|_{L^\infty(D)} < \infty,
\end{aligned}$$

which is bounded by the analyticity of the diffusion coefficient, i.e.,

$$\frac{a(\mathbf{z}_h) - a(\mathbf{z})}{h} \to \partial_n a(\mathbf{z}) \in L^\infty(D), \quad h \to 0.$$

The second term in (3.18) tends to 0 as well: The operator norm can be bounded by

$$\begin{aligned}
\left\| \frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} - \partial_n A_{\mathbf{z}} \right\| &= \sup_{u,v \in H_0^1(D), u,v \neq 0} \frac{\left| [\frac{A_{\mathbf{z}_h} - A_{\mathbf{z}}}{h} u](v) - [\partial_n A_{\mathbf{z}} u](v) \right|}{\|u\|_{H_0^1(D)} \|v\|_{H_0^1(D)}} \\
&\leq \sup_{u,v \in H_0^1(D), u,v \neq 0} \frac{\left| \int_D \frac{a(\mathbf{z}_h) - a(\mathbf{z})}{h} \nabla u \nabla v - \partial_n a(\mathbf{z}) \nabla u \nabla v \, \mathrm{d}\mathbf{x} \right|}{\|u\|_{H_0^1(D)} \|v\|_{H_0^1(D)}} \\
&= \sup_{u,v \in H_0^1(D), u,v \neq 0} \frac{\left| \int_D (\frac{a(\mathbf{z}_h) - a(\mathbf{z})}{h} - \partial_n a(\mathbf{z})) \nabla u \nabla v \, \mathrm{d}\mathbf{x} \right|}{\|u\|_{H_0^1(D)} \|v\|_{H_0^1(D)}} \\
&\leq \left\| \frac{a(\mathbf{z}_h) - a(\mathbf{z})}{h} - \partial_n a(\mathbf{z}) \right\|_{L^\infty(D)} \to 0, \quad h \to 0.
\end{aligned}$$

$\|u(\mathbf{z})\|_{H_0^1(D)}$ is bounded by

$$\|u(\mathbf{z})\|_{H_0^1(D)} \leq \frac{C_P}{a_{min}(\mathbf{z})} \|f(\mathbf{z})\|_{L^2(D)} < \infty$$

by the assumptions on the data. Thus, it is $L_h \to L_0$ as $h \to 0$.

Step 2d: Show that the partial derivative $\partial_n u(\mathbf{z})$ exists.
For the right-hand side of $A_{\mathbf{z}} q_h = L_h$ it is $L_h \to L_0$ as $h \to 0$. Since q_h depends continuously on L_h, it is $q_h \to q_0 = \partial_n u(\mathbf{z})$ as $h \to 0$ with

$$A_{\mathbf{z}} \partial_n u = L_0.$$

Therefore, the partial derivative $\partial_n u(\mathbf{z})$ of $u(\mathbf{z})$ exists. ■

Remark 3.58 – Proof for the mixed form:
The proof for the mixed form is similar to the proof given above. The first part, i.e., $\mathrm{int}(\Sigma_U)$ is open, is the same. The second part follows the same steps as the proof for the diffusion problem and is similar to the proof given in [11, pp. 7–8] (only on the space $\mathrm{int}(\Sigma_U)$ instead of Σ_D), and is therefore omitted here. □

Remark 3.59 – Relevance of the solution's continuity:
In the proof, the continuity of the solution in $\mathrm{int}(\Sigma_U)$ was used. The continuity was shown in Lemma 3.25 based on Assumption 3.24 on the data in $\mathbb{C}^N$.

A similar result – without assuming a continuous solution – can be obtained by applying the Theorem of Hartogs (see Theorem C.4). This theorem ensures the continuity of the solution u in $\mathbf{z} \in \text{int}(\Sigma_U)$ if the analyticity of $z_n \mapsto u(z_1, \ldots, z_{n-1}, z_n, z_{n+1}, \ldots, z_N)$ can be verified.
If the existence of $\lim_{h\to 0} q_h(\mathbf{z})$ can be shown, differentiability in direction n is ensured, and with the Theorem of Osgood (see Theorem C.2), applied to a function in one variable, holomorphy in this direction. This means, $z_n \mapsto u(z_1, \ldots, z_{n-1}, z_n, z_{n+1}, \ldots, z_N)$ is analytic. Hence, by the Theorem of Hartogs u is continuous in $\text{int}(\Sigma_U)$. Since u is also partial differentiable, it is weakly holomorphic and by the Theorem of Osgood analytic.
This strategy was applied in [8, p. 10]. □

3.4.2.2 ***D*-bounded diffusion coefficient** In the case of a D-bounded diffusion coefficient similar assumptions on the analyticity of the data as in the bounded case have to be stated.

Assumption 3.60 – Analytic data – *D*-bounded coefficient:
Let the functions $b : \Sigma_D \to L^\infty(D)$ and $f : \Sigma_D \to L^2(D)$ be analytic in Σ_D. □

Theorem 3.61 – Analyticity of w in product domain of $\mathbb{C}^N$:
Let Assumption 3.60 and the assumptions of Lemma 3.38 and 3.41 be fulfilled, respectively, i.e., there exists a unique solution $\mathbf{w} \in C^0_\sigma(\Sigma_D; W(D))$.
Further, let there exist a unique solution $\mathbf{w} \in L^2_\rho(\Gamma; W(D))$, *which is fulfilled by the assumptions of Lemma 3.46 and 3.48, respectively.*
Then, also the mapping $\mathbf{z} \mapsto \mathbf{w}(\mathbf{z})$ *is analytic in* Σ_D. *Hence the solution* $\mathbf{w} \in L^2_\rho(\Gamma; W(D))$, $\mathbf{w} : \Gamma \to W(D)$ *can be analytically extended to* Σ_D. □

Proof: The proof is similar to the proof of Theorem 3.57. It consists in showing that the set Σ_D is an open set and that the map $\mathbf{z} \mapsto \mathbf{w}(\mathbf{z})$ is analytic in Σ_D.
First, it will be shown that Σ_D is an open set. By Definition 3.33 for $\bar{a}_{min}(\mathbf{z})$ and $\bar{a}_{max}(\mathbf{z})$ the functions $\bar{a}_{min}(\mathbf{z})$ and $\bar{a}_{max}(\mathbf{z})$ are continuous in $\mathbf{z} \in \mathbb{C}^N$. By the property that a function from $X \to Y$ is continuous if and only if the preimage of open sets in Y are open sets in X, the functions $\bar{a}_{min}(\mathbf{z})$ as well as $\bar{a}_{max}(\mathbf{z})$ are continuous functions in $\mathbf{z} \in \mathbb{C}^N$. The set $(0, \infty)$ is open in $\mathbb{R}$, and so is the set $\{\mathbf{z} \in \mathbb{C}^N : 0 < \bar{a}_{min}(\mathbf{z})\}$. Since $(-\infty, \infty)$ is open in $\mathbb{R}$, the set $\{\mathbf{z} \in \mathbb{C}^N : \bar{a}_{max}(\mathbf{z}) < \infty\}$ is open. The intersection of the open sets $\{\mathbf{z} \in \mathbb{C}^N : 0 < \bar{a}_{min}(\mathbf{z})\}$ and $\{\mathbf{z} \in \mathbb{C}^N : \bar{a}_{max}(\mathbf{z}) < \infty\}$ is again open (while non-empty because $\bar{a}_{min}(\mathbf{z}) \leq \bar{a}_{max}(\mathbf{z})$ by assumption) and therefore,

$$\begin{aligned}\Sigma_D &= \{\mathbf{z} \in \mathbb{C}^N : 0 < \bar{a}_{min}(\mathbf{z})\} \cap \{\mathbf{z} \in \mathbb{C}^N : \bar{a}_{max}(\mathbf{z}) < \infty\} \\ &= \{\mathbf{z} \in \mathbb{C}^N : 0 < \bar{a}_{min}(\mathbf{z}) \leq \bar{a}_{max}(\mathbf{z}) < \infty\}\end{aligned}$$

is an open set.

The analyticity of the diffusion problem can be shown as before by substituting Σ_D for Σ_U. The proof for the mixed form has a similar structure and can be found in [11, pp. 7–8]. ■

Remark 3.62:
All the results given in this section could have been shown in the same way for f in $H^{-1}(D)$, the dual space of $H_0^1(D)$. □

3.5 Additional regularity results

In order to derive error estimates of the spatial discretization and in particular of the multilevel method (see Section 5), where spatial and stochastic variables are coupled, additional results on the regularity of the solution are required.

3.5.1 Diffusion problem

In order to derive H^2-regularity of the weak solution of (3.5) with a uniformly bounded diffusion coefficient the following assumption needs to be fulfilled.

Assumption 3.63 – Data to derive $u \in H^2(D)$, a uniformly bdd.:
Assume that $f \in L_\rho^r(\Gamma; L^2(D))$ for some $2 \leq r < \infty$ and $a(\mathbf{y}) \in W^{1,\infty}(D)$ (see Definition B.3) for every $\mathbf{y} \in \Gamma$. Further, let $\nabla a \in L^\infty(\Gamma; L^\infty(D))$. □

Before stating the regularity result, for simplicity of notation, introduce the space $\mathcal{W}$ defined as

$$\mathcal{W} = H^2(D) \cap H_0^1(D)$$

equipped with norm

$$\|u\|_{\mathcal{W}} = \|u\|_{H_0^1(D)} + \|\Delta u\|_{L^2(D)}.$$

The following theorem can be found in [2, pp. 129 – 130], where only $\|a\|_{L^\infty(\Gamma;W^{1,\infty}(D))}$ instead of $\|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))}$ is used.

Theorem 3.64 – Solution in $H^2(D)$, uniformly bounded a:
Let Assumption 3.63 hold. Then, the weak solution u of the diffusion problem (3.5) with uniformly bounded diffusion coefficient (i.e., the assumptions of Theorem 3.21 should hold) is for all $\mathbf{y} \in \Gamma$ in $H^2(D)$ and hence in $\mathcal{W}$. Further, it is $u \in L_\rho^r(\Gamma; \mathcal{W})$.

For $2 \leq r < \infty$ as in Assumption 3.63 the following estimate can be derived

$$\|u\|_{L^r_\rho(\Gamma;\mathcal{W})} \leq C(a)\|f\|_{L^r_\rho(\Gamma;L^2(D))},$$

where $C(a) > 0$ is a constant only depending on a_{min} and $\|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))}$, but not on $\mathbf{y} \in \Gamma$. □

Proof: The sketch of the proof in [2, pp. 129 − 130] is carried out in more detail. It holds for all $\mathbf{y} \in \Gamma$, $\mathbf{x} \in D$

$$f(\mathbf{y},\mathbf{x}) = -\nabla \cdot (a(\mathbf{y},\mathbf{x})\nabla u(\mathbf{y},\mathbf{x})) = -\nabla a(\mathbf{y},\mathbf{x}) \cdot \nabla u(\mathbf{y},\mathbf{x}) - a(\mathbf{y},\mathbf{x})\Delta u(\mathbf{y},\mathbf{x})$$

and hence,

$$\begin{aligned} f(\mathbf{y},\mathbf{x}) + \nabla a(\mathbf{y},\mathbf{x}) \cdot \nabla u(\mathbf{y},\mathbf{x}) &= -a(\mathbf{y},\mathbf{x})\Delta u(\mathbf{y},\mathbf{x}) \\ \Leftrightarrow \frac{1}{a(\mathbf{y},\mathbf{x})}\big(f(\mathbf{y},\mathbf{x}) + \nabla a(\mathbf{y},\mathbf{x}) \cdot \nabla u(\mathbf{y},\mathbf{x})\big) &= -\Delta u(\mathbf{y},\mathbf{x}). \end{aligned}$$

By this equation Δu exists and

$$\begin{aligned} \|\Delta u(\mathbf{y},\cdot)\|_{L^2(D)} &= \left\| \frac{1}{a(\mathbf{y},\cdot)}\big(f(\mathbf{y},\cdot) + \nabla a(\mathbf{y},\cdot) \cdot \nabla u(\mathbf{y},\cdot)\big) \right\|_{L^2(D)} \\ &\leq \left\| \frac{1}{a(\mathbf{y},\cdot)} \right\|_{L^\infty(D)} \|f(\mathbf{y},\cdot) + \nabla a(\mathbf{y},\cdot) \cdot \nabla u(\mathbf{y},\cdot)\|_{L^2(D)} \\ &\leq \frac{1}{a_{min}}\big(\|f(\mathbf{y},\cdot)\|_{L^2(D)} + \|\nabla a(\mathbf{y},\cdot) \cdot \nabla u(\mathbf{y},\cdot)\|_{L^2(D)}\big). \end{aligned}$$

Using equation (3.6) for $\Gamma \subset \mathbb{R}^N$ it holds

$$\begin{aligned} \|\nabla a(\mathbf{y},\cdot) \cdot \nabla u(\mathbf{y},\cdot)\|_{L^2(D)} &\leq \|\nabla a(\mathbf{y},\cdot)\|_{L^\infty(D)} \|u(\mathbf{y},\cdot)\|_{H^1_0(D)} \\ &\overset{(3.6)}{\leq} \|\nabla a(\mathbf{y},\cdot)\|_{L^\infty(D)} \frac{C_P}{a_{min}} \|f(\mathbf{y},\cdot)\|_{L^2(D)} \end{aligned}$$

and it follows

$$\|\Delta u(\mathbf{y},\cdot)\|_{L^2(D)} \leq \frac{a_{min} + C_P\|\nabla a(\mathbf{y},\cdot)\|_{L^\infty(D)}}{a^2_{min}} \|f(\mathbf{y},\cdot)\|_{L^2(D)} \tag{3.19}$$

which is finite for each $\mathbf{y} \in \Gamma$ because $\nabla a(\mathbf{y},\cdot) \in L^\infty(D)$ and $f(\mathbf{y},\cdot) \in L^2(D)$. $u(\mathbf{y},\cdot) \in H^1_0(D)$ has already been shown in Theorem 3.21. Thus, $u(\mathbf{y},\cdot) \in \mathcal{W}$.

The second statement of the theorem is obtained through the definition of the norm $\|\cdot\|_{\mathcal{W}}$, the stability estimate (3.6) of the solution $\|u(\mathbf{y},\cdot)\|^r_{H^1_0(D)}$ given in Theorem 3.21, equation (3.19) with

$$\|\nabla a(\mathbf{y},\cdot)\|_{L^\infty(D)} \leq \|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))} \tag{3.20}$$

and the assumptions $f \in L^r_\rho(\Gamma;L^2(D))$ and $\nabla a \in L^\infty(\Gamma;L^\infty(D))$

$$\begin{aligned}
&\|u\|^r_{L^r_\rho(\Gamma;\mathcal{W})} \\
&= \int_\Gamma \|u(\mathbf{y},\cdot)\|^r_{\mathcal{W}} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \int_\Gamma \left(\|u(\mathbf{y},\cdot)\|_{H^1_0(D)} + \|\Delta u(\mathbf{y},\cdot)\|_{L^2(D)}\right)^r \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{(3.6)}{\leq} \int_\Gamma \left(\frac{C_P}{a_{min}}\|f(\mathbf{y},\cdot)\|_{L^2(D)} + \|\Delta u(\mathbf{y},\cdot)\|_{L^2(D)}\right)^r \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{(3.19)}{\leq} \int_\Gamma \left(\frac{C_P}{a_{min}} + \frac{a_{min} + C_P\|\nabla a(\mathbf{y},\cdot)\|_{L^\infty(D)}}{a^2_{min}}\right)^r \|f(\mathbf{y},\cdot)\|^r_{L^2(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{(3.20)}{\leq} \left(\frac{C_P}{a_{min}} + \frac{a_{min} + C_P\|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))}}{a^2_{min}}\right)^r \int_\Gamma \|f(\mathbf{y},\cdot)\|^r_{L^2(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\leq C(a)^r \|f\|^r_{L^r_\rho(\Gamma;L^2(D))} < \infty.
\end{aligned}$$

■

The next aim is to derive a similar result for the diffusion problem with a D-bounded diffusion coefficient. As the lower and upper bound of the coefficient depend on the stochastic variable $\mathbf{y} \in \Gamma$, slightly different assumptions on the data are required.

Assumption 3.65 – Data to derive $u \in H^2(D)$, a D-bounded:
Assume that $f \in L^s_\rho(\Gamma;L^2(D))$ for some $1 < s < \infty$, $a(\mathbf{y}) \in W^{1,\infty}(D)$ for every $\mathbf{y} \in \Gamma$ and $\nabla a \in L^\infty(\Gamma;L^\infty(D))$. Further, let $\frac{1}{a_{min}} \in L^q_\rho(\Gamma;\mathbb{R})$ for some $q \in [1,\infty)$ chosen according to s. □

Theorem 3.66 – Solution in $H^2(D)$, a D-bounded:
Let Assumption 3.65 on the data hold. Then, with the assumptions of Theorem 3.36 the weak solution u of the diffusion problem (3.5) with D-bounded diffusion coefficient is for all $\mathbf{y} \in \Gamma$ in $H^2(D)$ and hence in $\mathcal{W}$.
Further, it is $u \in L^r_\rho(\Gamma;\mathcal{W})$ for $2 \leq r < \infty$.
For $2 \leq r < \infty$ and $1 < s < \infty$ the following estimate can be derived

$$\|u\|_{L^r_\rho(\Gamma;\mathcal{W})} \leq C(a)\|f\|_{L^{sr}_\rho(\Gamma;L^2(D))},$$

where $C(a)$ is a constant independent of $\mathbf{y} \in \Gamma$, depending on $\|\frac{1}{a_{min}}\|_{L^{q^*}_\rho(\Gamma;\mathbb{R})}$ for values of q^* in $[1,\infty)$, $\|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))}$ and r. □

Proof: The first part of the proof is similar to the proof of the previous theorem, yielding

$$\|\Delta u(\mathbf{y},\cdot)\|_{L^2(D)} \leq \frac{a_{min}(\mathbf{y}) + C_P\|\nabla a(\mathbf{y},\cdot)\|_{L^\infty(D)}}{a^2_{min}(\mathbf{y})}\|f(\mathbf{y},\cdot)\|_{L^2(D)}, \qquad (3.21)$$

which is finite for each $\mathbf{y} \in \Gamma$ because $\nabla a(\mathbf{y},\cdot) \in L^\infty(D)$, $a_{min}(\mathbf{y}) > 0$ and $f(\mathbf{y},\cdot) \in L^2(D)$. Due to Theorem 3.36 it is $u(\mathbf{y},\cdot) \in H^1_0(D)$. Hence, $u(\mathbf{y},\cdot) \in \mathcal{W}$ for all $\mathbf{y} \in \Gamma$.

The second part of the theorem follows by equation (3.21), the stability estimate on the solution derived in Theorem 3.36 as well as the Hölder inequality (see Lemma B.2) with $s, q > 1$ and $\frac{1}{s} + \frac{1}{q} = 1$

$$\begin{aligned}
&\|u\|^r_{L^r_\rho(\Gamma;\mathcal{W})} \\
&\quad\leq \int_\Gamma \left(\|u(\mathbf{y},\cdot)\|_{H^1_0(D)} + \|\Delta u(\mathbf{y},\cdot)\|_{L^2(D)}\right)^r \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\quad\overset{(3.21)}{\leq} \int_\Gamma \left(\frac{C_P}{a_{min}(\mathbf{y})} + \frac{a_{min}(\mathbf{y}) + C_P\|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))}}{a^2_{min}(\mathbf{y})}\right)^r \|f(\mathbf{y},\cdot)\|^r_{L^2(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\quad= \int_\Gamma c^r(a)\|f(\mathbf{y},\cdot)\|^r_{L^2(D)}\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\quad\overset{\text{Lemma B.2}}{\leq} \left(\int_\Gamma c^{rq}(a)\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\right)^{1/q} \cdot \|f\|^r_{L^{rs}_\rho(\Gamma;L^2(D))}
\end{aligned}$$

which gives the result by defining $C(a) = \left(\int_\Gamma c^{rq}(a)\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\right)^{1/rq}$, where the constant $C(a)$ depends on $\frac{1}{a_{min}(\mathbf{y})}$, $\|\nabla a\|_{L^\infty(\Gamma;L^\infty(D))}$ and r. ■

3.5.2 Mixed form of the diffusion problem

Now, conditions on p are stated in the same way as in the previous subsection. In Section 5 the solution $(p, \mathbf{u})$ will be assumed to be in the right space because for $\mathbf{u}$ no references concerning the regularity are known to the author.

The following theorem and the proof are stated similarly to the one for the diffusion problem, where $H^1(D)$-regularity (instead of $H^2(D)$-regularity) is required in Section 5 for p.

Theorem 3.67 – p in $H^1(D)$, a uniformly bdd.:
Let $f \in L^r_\rho(\Gamma; L^2(D))$ for some $2 \leq r < \infty$ and the assumptions of Theorem 3.29 on Γ hold.

Then, the part of the unique weak solution p of the diffusion problem's mixed form (3.9) with uniformly bounded diffusion coefficient is for all $\mathbf{y} \in \Gamma$ in $H^1(D)$. Further, it is $p \in L^r_\rho(\Gamma; H^1(D))$. $\square$

Proof: In a first step, it has to be shown that for all $\mathbf{y} \in \Gamma$ it is $p(\mathbf{y}) \in H^1(D)$. Since $p(\mathbf{y}) \in L^2(D)$ by Theorem 3.29, it suffices to show $\|\nabla p(\mathbf{y})\|_{L^2(D)} < \infty$. It is by the first line of (3.2) on page 14 and $\|\frac{1}{a(\mathbf{y})}\|_{L^\infty(D)} \leq \frac{1}{a_{min}}$

$$\|\nabla p(\mathbf{y})\|_{L^2(D)} = \left\| \frac{1}{a(\mathbf{y})} \mathbf{u}(\mathbf{y}) \right\|_{L^2(D)} \leq \frac{1}{a_{min}} \|\mathbf{u}(\mathbf{y})\|_{L^2(D)}. \tag{3.22}$$

Since $\mathbf{u}(\mathbf{y}) \in H(\mathrm{div}; D)$, it is in particular $\mathbf{u}(\mathbf{y}) \in [L^2(D)]^d$ by the definition of $H(\mathrm{div}; D)$. Thus, it holds $\|\mathbf{u}(\mathbf{y})\|_{L^2(D)} < \infty$ and $p(\mathbf{y}) \in H^1(D)$.

To derive the second statement, i.e., $p \in L^r_\rho(\Gamma; H^1(D))$, the crucial step is to use the estimates given in Theorem 3.29. Further, the assumption on f and the just given estimates on the norm have to be applied. Note that

$$\begin{aligned}
&\|p(\mathbf{y})\|^2_{L^2(D)} + \|\nabla p(\mathbf{y})\|^2_{L^2(D)} \\
\overset{(3.22)}{\leq}\; &\|p(\mathbf{y})\|^2_{L^2(D)} + C\left(\frac{1}{a_{min}}\right)^2 \|\mathbf{u}\|^2_{L^2(D)} \\
\overset{\text{Thm. 3.29}}{\leq}\; &C\left(\frac{a_{max}}{a_{min}^2}\right)^2 \|f(\mathbf{y})\|^2_{L^2(D)} + C\left(\frac{1}{a_{min}}\right)^2 \left(\frac{a_{max}}{a_{min}}\right)^2 \|f(\mathbf{y})\|^2_{L^2(D)},
\end{aligned}$$

where $\|\mathbf{u}(\mathbf{y})\|_{L^2(D)} \leq \|\mathbf{u}(\mathbf{y})\|_{H(\mathrm{div};D)}$ has been used from the second to the last line to apply the first estimate of Theorem 3.29.
With the previous estimate, the estimate for $p \in L^r_\rho(\Gamma; H^1(D))$ follows by

$$\begin{aligned}
\|p\|^r_{L^r_\rho(\Gamma; H^1(D))} &= \int_\Gamma \|p(\mathbf{y})\|^r_{H^1(D)} \rho(\mathbf{y})\, \mathrm{d}\mathbf{y} \\
&= \int_\Gamma \left(\|p(\mathbf{y})\|^2_{L^2(D)} + \|\nabla p(\mathbf{y})\|^2_{L^2(D)} \right)^{r/2} \rho(\mathbf{y})\, \mathrm{d}\mathbf{y} \\
&\leq \int_\Gamma C^r \left(\frac{a_{max}}{a_{min}^2}\right)^r \|f(\mathbf{y})\|^r_{L^2(D)} \rho(\mathbf{y})\, \mathrm{d}\mathbf{y} \\
&= C^r \left(\frac{a_{max}}{a_{min}^2}\right)^r \|f\|^r_{L^r_\rho(\Gamma; L^2(D))} < \infty.
\end{aligned}$$

Hence, $p \in L^r_\rho(\Gamma; H^1(D))$. $\blacksquare$

Assumption 3.68 – For additional regularity, a D-bounded:
Let $2 \leq r < \infty$. Assume for the $L^r_\rho(\Gamma; H^1(D))$-regularity of p that for $q > 1$ it holds $f \in L^{rq}_\rho(\Gamma; L^2(D))$, for some $s > 2r$ it holds $\frac{1}{a_{min}} \in L^s(\Gamma; \mathbb{R})$ and let $a_{max} \in L^{s^*}_\rho(\Gamma; L^\infty(D))$ for $s^* > 1$ chosen according to s. $\square$

Remark 3.69:
The previous assumption can be specified when it is clear which specific regularity of the solution p is necessary. For an increasing number of r the Hölder inequality has to be applied several times. This determines the values of s^*. The previous assumption is stated in a more general way in order to reduce the notation. □

Theorem 3.70 – p in $H^1(D)$, D-bounded a:
Let Assumption 3.68 as well as the assumptions of Theorem 3.39 on the existence of a unique solution $(p(\mathbf{y}), \mathbf{u}(\mathbf{y}))$ for $\mathbf{y} \in \Gamma$ hold.
Then, $p(\mathbf{y})$ of the diffusion problem's mixed form (3.9) with D-bounded diffusion coefficient is for all $\mathbf{y} \in \Gamma$ in $H^1(D)$. Further, it is $p \in L^r_\rho(\Gamma; H^1(D))$. □

Proof: The first part of the theorem, i.e., $p(\mathbf{y})$ in $H^1(D)$ for all $\mathbf{y} \in \Gamma$ follows straightforward from the proof with a uniformly bounded diffusion coefficient by equation (3.22) adjusted for a D-bounded diffusion coefficient, i.e.,

$$\|\nabla p(\mathbf{y})\|_{L^2(D)} \leq \frac{1}{a_{min}(\mathbf{y})} \|\mathbf{u}(\mathbf{y})\|_{L^2(D)}. \tag{3.23}$$

The assumption on the data give the result.

For the second part of the theorem $\|p\|_{L^r_\rho(\Gamma;H^1(D))}$ has to be bounded. This is done by applying the same estimates as in the proof for the uniformly bounded diffusion coefficient – but now the coefficient depends on $\mathbf{y}$ – on $\|p(\mathbf{y})\|_{H^1(D)}$ and by applying the Hölder inequality (see Lemma B.2).

$$\begin{aligned}
\|p\|^r_{L^r_\rho(\Gamma;H^1(D))} &= \int_\Gamma \|p(\mathbf{y})\|^r_{H^1(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{\text{prev. proof}}{\leq} \int_\Gamma C^r \left(\frac{a_{max}(\mathbf{y})}{a^2_{min}(\mathbf{y})}\right)^r \|f(\mathbf{y})\|^r_{L^2(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{\text{Lemma B.2}}{\leq} C^r \left(\int_\Gamma \left(\frac{a_{max}(\mathbf{y})}{a^2_{min}(\mathbf{y}}\right)^{rs} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\right)^{1/s} \|f\|^r_{L^{rq}(\Gamma;L^2(D))}.
\end{aligned}$$

With a Hölder inqequality on $\int_\Gamma \left(\frac{a_{max}(\mathbf{y})}{a^2_{min}(\mathbf{y}}\right)^{rs} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y}$ the result follows. ■

4 Single-level sparse grid approximations and quadrature

So far, the regularity of the diffusion problem's and its mixed form's solution have been discussed. This section's content is how such a solution – as no closed form is known – can be numerically approximated. As mentioned in the introduction, the solution might be approximated by applying a spectral stochastic Galerkin method or stochastic collocation. Also, an approximation by a Monte-Carlo method is possible.

The focus in this thesis is on the collocation method. Stochastic Galerkin and stochastic collocation methods differ in the coupling of the spatial and stochastic variable. In stochastic Galerkin approximations the variables are coupled, while they are uncoupled by construction in the stochastic collocation method (see below). The collocation method is a sampling method as Monte-Carlo methods. In contrast to them, in stochastic collocation methods the stochastic problem is transformed into a deterministic problem. This is done by choosing a deterministic grid as the required point set instead of generating the point set from a certain probability distribution. The approximations in the spatial variable are collocated in the chosen point set. Therefore, the method is called collocation method. The collocation leads to an uncoupling of spatial and stochastic variables. Different kinds of grids, like tensor product grids or sparse grids, can be used in the stochastic collocation method for approximating the solution. An approximation with a full tensor product grid is given in [1]. Since the computational complexity of full tensor product approximations grows exponentially with increasing number of dimensions N of Γ, approximations on sparse grids are often applied. In [22] and [11] isotropic Smolyak sparse grids are considered, while anisotropic ones in [23]. Isotropic means that all directions are weighted equally, while in anisotropic grids the directions are weighted differently. The stochastic collocation method beats the Monte-Carlo method for a moderate number of dimensions N (see [22, pp. 2339 – 2340]). The construction of so-called quasi-optimal grids is given in [4] and [21].

Quadrature is of interest when determining the mean value, variance or higher moments of the solution. A quadrature method can also be considered on tensor product and sparse grids. If the deterministic grid is chosen in a certain way, the expected value of the approximated solution obtained by the collocation method turns out to be equivalent to directly applying quadrature for obtaining the solution's expected value.

The section shortly presents the collocation method and quadrature on tensor product and sparse grids. The connection of numeric quadrature to stochastic collocation is pointed out, and a choice of abscissae of the deterministic grid is given to derive the equivalence of both methods when the expected value of the solution is the value of interest. The section is concluded by the error estimates on the tensor product and Smolyak sparse grid approximations, and an error estimate on the tensor product quadrature formula.

4.1 Approximations by Stochastic Collocation

In the sequel, the collocation method used to solve a stochastic elliptic boundary value problem is presented. It gives insight why the fully weak, i.e., with respect to spatial and stochastic domain, or just the weak solution, i.e., the solution with respect to the spatial domain, can be considered in the analysis when deriving results of the existence and uniqueness of a solution.

It is proceeded as follows: A discretization of the stochastic and spatial domain is given, the semidiscrete approximation is defined, and then, the method itself is specified. The latter consists of 3 steps:

1. selection of a point set,
2. computation of the semidiscrete solution in the just selected point set,
3. approximation of the solution.

In order to compute the weak solution (2.4) a discretization of the function space is necessary. In the following, the space $L^q_\rho(\Gamma; W(D))$ is assumed to factorize in $L^q_\rho(\Gamma) \otimes W(D)$, which is the case for the spaces mentioned in Remark 2.17. Consequently, the spatial and the parameter space can be discretized by their own. Denote

- $\mathcal{T}_h$ – regular triangulation of D with maximal mesh-spacing $h > 0$,
- $W_h(D) \subset W(D)$ – conforming finite element space of dimension N_h,
- $\{\boldsymbol{\phi}_i(\mathbf{x})\}_{i=1}^{N_h}$ – basis of $W_h(D)$.

For the parameter domain Γ denote

- $\mathcal{P}(\Gamma) \subset L^q_\rho(\Gamma)$ – polynomial subspace of finite dimension η,
- $\{\psi_i(\mathbf{y})\}_{i=1}^{\eta}$ – basis of $\mathcal{P}(\Gamma)$.

Recall the formulation of the weak solution: Find $\mathbf{w}(\mathbf{y},\mathbf{x}) \in L^q_\rho(\Gamma) \otimes W(D)$ such that for all $\mathbf{v}(\mathbf{y},\mathbf{x}) \in L^q_\rho(\Gamma) \otimes W(D)$ it holds

$$\int_\Gamma \int_D \mathcal{L}\big(a(\mathbf{y},\mathbf{x})\big)\big(\mathbf{w}(\mathbf{y},\mathbf{x})\big)\mathbf{v}(\mathbf{y},\mathbf{x})\rho(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = \int_\Gamma \int_D \mathbf{f}(\mathbf{y},\mathbf{x})\mathbf{v}(\mathbf{y},\mathbf{x})\rho(\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}$$

with additional equations for suitable boundary conditions.
Since the space $L^q_\rho(\Gamma; W(D))$ factorizes in $L^q_\rho(\Gamma) \otimes W(D)$, an equivalent form is given by: Find $\mathbf{w}(\mathbf{x},\mathbf{y}) \in L^q_\rho(\Gamma) \otimes W(D)$ such that for all $\mathbf{v}(\mathbf{x}) \in W(D)$ it holds ρ-a.e. in Γ

$$\int_D \mathcal{L}\big(a(\mathbf{y},\mathbf{x})\big)\big(\mathbf{w}(\mathbf{y},\mathbf{x})\big)\mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_D \mathbf{f}(\mathbf{y},\mathbf{x})\mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{4.1}$$

with additional equations for suitable boundary conditions.
Now, using formulation (4.1), only the spatial variable can be discretized, and the so-called *semidiscrete approximation* $\mathbf{w}_h \in L^q_\rho(\Gamma) \otimes W_h(D)$ given by

$$\mathbf{w}_h(\mathbf{y},\mathbf{x}) = \sum_{j=1}^{N_h} c_j(\mathbf{y})\boldsymbol{\phi}_j(\mathbf{x}) \tag{4.2}$$

is obtained. The $c_j(\mathbf{y})$ are coefficients which can be determined by solving for all $j' \in \{1,\ldots,N_h\}$ the problem

$$\int_D \mathcal{L}\big(a(\mathbf{y},\mathbf{x})\big)\left(\sum_{j=1}^{N_h} c_j(\mathbf{y})\boldsymbol{\phi}_j(\mathbf{x})\right)\boldsymbol{\phi}_{j'}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_D \mathbf{f}(\mathbf{y},\mathbf{x})\boldsymbol{\phi}_{j'}(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{4.3}$$

Note that this system of equations consists of N_h equations to be solved.
The semidiscrete approximation (4.2) is used in the collocation method whose above mentioned steps are explicated. After

1. having selected a point set $\{\mathbf{y}_i : \mathbf{y}_i \in \Gamma \subset \mathbb{R}^N\}_{i=1}^{\eta}$,

2. the semidiscrete solution $\mathbf{w}_h(\mathbf{y}_i,\mathbf{x})$ is computed for all $\mathbf{y}_i$ in the selected set $\{\mathbf{y}_i : \mathbf{y}_i \in \Gamma \subset \mathbb{R}^N\}_{i=1}^{\eta}$, where $\mathbf{w}_h(\mathbf{y}_i,\mathbf{x}) = \sum_{j=1}^{N_h} c_{j,i}\boldsymbol{\phi}_j(\mathbf{x})$. The coefficients $c_{j,i} = c_j(\mathbf{y}_i)$ are obtained by solving the weak formulation (4.3), where the random variable $\mathbf{y}$ is constant, namely $\mathbf{y}_i$, such that

$$\int_D \mathcal{L}\big(a(\mathbf{y}_i,\mathbf{x})\big)\left(\sum_{j=1}^{N_h} c_{j,i}\boldsymbol{\phi}_j(\mathbf{x})\right)\boldsymbol{\phi}_{j'}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_D \mathbf{f}(\mathbf{y}_i,\mathbf{x})\boldsymbol{\phi}_{j'}(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{4.4}$$

$$\text{for } j' \in \{1,\ldots,N_h\}$$

with a boundary condition on $\sum_{j=1}^{N_h} c_{j,i}\boldsymbol{\phi}_j(\mathbf{x})$. Integrals in (4.4) are for instance computed by a suitable quadrature rule depending on the triangulation $\mathcal{T}_h$.

3., the solution is approximated by some approximation rule called $\mathcal{A}$, i.e., $\mathbf{w}_{h,\eta}(\mathbf{y},\mathbf{x}) = \mathcal{A}(\{\mathbf{w}_h(\mathbf{y}_i,\mathbf{x})\}_{i=1}^{\eta}, \{\psi_j(\mathbf{y})\}_{j=1}^{\eta})$.

The choice of the point set $\{\mathbf{y}_i : \mathbf{y}_i \in \Gamma \subset \mathbb{R}^N\}_{i=1}^{\eta}$ will be described in the following. It depends on the one hand on the selection of abscissae (see Section 4.1.1) and on the other hand on the kind of approximation $\mathcal{A}$, e.g., a full tensor product or a sparse grid approximation (see Sections 4.1.2 and 4.1.3).

The approximation of the solution might be performed by several different polynomial spaces and approximation rules applied to the semidiscrete solution. In the following, the tensor product grid and Smolyak sparse grid approximation based on Lagrangian polynomials will be specified.

Throughout the rest of the section, the semidiscrete approximation in the spatial variable, $\mathbf{w}_h(\mathbf{y},\mathbf{x}) = \sum_{j=1}^{N_h} c_j(\mathbf{y})\boldsymbol{\phi}_j(\mathbf{x})$, is assumed to be given. Thus, the interest is in the approximation in the parameter space which is approximated by global polynomials here.

4.1.1 Abscissae

Let $\mathbf{i} = (i_1, \dots, i_N)$ with $i_n \in \mathbb{N},\ i_n \geq 1$ for all $n \in \{1, \dots N\}$, be a multi-index. This index is introduced here because it will be needed later on in the definition of the sparse grid approximation.
Fix $n \in \{1, \dots, N\}$, and denote the abscissae in direction n related to the index i_n by

$$\theta^{i_n} = \{\theta_1^{i_n}, \theta_2^{i_n}, \dots, \theta_{m(i_n)}^{i_n}\} \subset \Gamma_n, \tag{4.5}$$

where $m : \mathbb{N} \to \mathbb{N}$ is a "level-to-nodes function", strictly increasing with $m(0) = 0, m(1) = 1$, $m(i_n) < m(i_n + 1)$, with $m(i_n)$ denoting the number of collocation points for the i_n-th index on which the interpolant will be defined later.
The abscissae $\{\theta_1^{i_n}, \dots, \theta_{m(i_n)}^{i_n}\}$ have to be determined for a concrete choice of the collocation points, i.e., the point sets $\{\mathbf{y}_i : \mathbf{y}_i \in \Gamma \subset \mathbb{R}^N\}_{i=1}^{\eta}$. Here, possible choices are presented. Principally, nested and non-nested point sets are distinguished. A point set is called nested if $\theta^{i_n} \subset \theta^{i_n+1}$, and non-nested

otherwise. The points themselves are chosen depending on the parameter space Γ and the probability measure $\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}$ on the parameter space.
For the parameter spaces being relevant here, i.e., $[-1,1]^N$ and $(-\infty,\infty)^N$, the following choices can be mostly found in the literature: For the bounded space $[-1,1]^N$ with uniform measure the abscissae are either so-called Gauss-Legendre or so-called Clenshaw-Curtis points. In the case of the unbounded parameter space $(-\infty,\infty)^N$ with Gaussian measure the abscissae are so-called Gauss-Hermite points.
The *Clenshaw-Curtis points* are the extrema of the Chebychev polynomials, and they are given by

$$\theta_k^{i_n} = -\cos\left(\frac{\pi(k-1)}{m(i_n)-1}\right), \quad k \in \{1,\ldots,m(i_n)\}.$$

For the following choice, they are even nested: If $m(i_n)=1$, let $\theta_1^{i_n}=0$, and define the function m as

$$m(1)=1 \quad \text{and} \quad m(i_n)=2^{i_n-1}+1, \quad i_n>1. \tag{4.6}$$

The *Gaussian points* are the zeros of polynomials q in $\mathcal{P}_{m(i_n)}$ which are orthogonal to $\mathcal{P}_{m(i_n)-1}$ with respect to the density ρ_n in $L^2_{\rho_n}$ for each direction $n\in\{1,\ldots,N\}$ (recall that the density function ρ is assumed to factorize), i.e., these are all zeros of q with

$$0=\int_{\Gamma_n} q(y)r(y)\rho_n(y)\,\mathrm{d}y \quad \forall\, r\in\mathcal{P}_{m(i_n)-1}.$$

For the uniform density Legendre polynomials are chosen. The points are called Gauss-Legendre points. For the Gaussian density and Hermite polynomials they are called Gauss-Hermite points. The resulting point sets are in general non-nested.

In Figure 1 on page 58 in the first column and the first row for the case $N=2$ the Clenshaw-Curtis abscissae in direction 1 and 2 can be found for $i_1,i_2\in\{1,\ldots,4\}$. The number of points $m(i_n)$ is determined by the rule given in (4.6). The Clenshaw-Curtis points are chosen because the idea is the same as for Gaussian points, but the implementation is more straight forward.

4.1.2 Tensor product approximation

By choosing different basis functions of the polynomial space, different collocation methods can be constructed. Here, a Lagrange basis is considered.

The Lagrange basis of the space $\mathcal{P}_{m(i_n)-1}(\Gamma_n)$ is given by $\{\lambda_j^{i_n}(y)\}_{j=1}^{m(i_n)}$, where

$$\lambda_j^{i_n}(y) = \prod_{m=1, m\neq j}^{m(i_n)} \frac{y - \theta_m^{i_n}}{\theta_j^{i_n} - \theta_m^{i_n}}$$

with the $\theta_m^{i_n}$ chosen according to the parameter space as described in the previous subsection.
Let

$$V_{m(i_n)}(\Gamma_n; W(D)) = \Big\{ \mathbf{v} \in C^0_\sigma(\Gamma_n; W(D)) : \mathbf{v}(y, \mathbf{x}) = \sum_{k=1}^{m(i_n)} \tilde{\mathbf{v}}_k(\mathbf{x}) \lambda_k^{i_n}(y),\ \{\tilde{\mathbf{v}}_k\}_{k=1}^{m(i_n)} \in W(D) \Big\}. \tag{4.7}$$

Further, denote by $\mathcal{U}^{i_n}$ the one-dimensional Lagrange interpolation operator related to index i_n with

$$\mathcal{U}^{i_n} : C^0_\sigma(\Gamma_n; W(D)) \to V_{m(i_n)}(\Gamma_n; W(D))$$

and

$$\mathcal{U}^{i_n}[\mathbf{w}_h](y) = \sum_{j=1}^{m(i_n)} \mathbf{w}_h(\theta_j^{i_n}) \lambda_j^{i_n}(y) \quad \forall\, \mathbf{w}_h \in C^0_\sigma(\Gamma_n; W(D)). \tag{4.8}$$

It is well known that this interpolation is exact for all polynomials of degree less than $m(i_n)$.

Let $\mathbf{i}$ be a fixed multi-index. The tensor product approximation based on Lagrangian polynomials, where $\mathbf{w}_h \in C^0_\sigma(\Gamma; W(D))$, is defined by

$$\begin{aligned} \mathcal{A}_N^{\mathbf{i}}[\mathbf{w}_h](\mathbf{y}) &= (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})[\mathbf{w}_h](\mathbf{y}) \\ &= \sum_{j_1=1}^{m(i_1)} \cdots \sum_{j_N=1}^{m(i_N)} \mathbf{w}_h(\theta_{j_1}^{i_1}, \ldots, \theta_{j_N}^{i_N})(\lambda_{j_1}^{i_1} \otimes \cdots \otimes \lambda_{j_N}^{i_N})(\mathbf{y}). \end{aligned}$$

In order to simplify notation let

$$J = \{1, \ldots, m(i_1)\} \times \cdots \times \{1, \ldots, m(i_N)\}$$

be the set of all possible combinations of indices $j_1, \ldots, j_N$.
Let for all $\mathbf{j}$ in J

$$\mathbf{y_j} = (\theta_{j_1}^{i_1}, \ldots, \theta_{j_N}^{i_N}).$$

With this definition the collocation point set is given by

$$\{\mathbf{y}_i : \mathbf{y}_i \in \Gamma \subset \mathbb{R}^N\}_{i=1}^{\eta} = \{\mathbf{y}_{\mathbf{j}}\}_{\mathbf{j}\in J} = (\theta^{i_1} \times \cdots \times \theta^{i_N}),$$

where for the dimension η it holds

$$\eta = |J| = \prod_{n=1}^{N} m(i_n).$$

Further, let

$$\lambda_{\mathbf{j}}(\mathbf{y}) = (\lambda_{j_1}^{i_1} \otimes \cdots \otimes \lambda_{j_N}^{i_N})(\mathbf{y}) = \prod_{k=1}^{N} \lambda_{j_k}^{i_k}(y_{i_k}).$$

The basis $\{\psi_i(\mathbf{y})\}_{i=1}^{\eta}$ which was defined in Section 4.1 is given by

$$\{\psi_i(\mathbf{y})\}_{i=1}^{\eta} = \{\lambda_{\mathbf{j}}(\mathbf{y})\}_{\mathbf{j}\in J}.$$

With this, the above tensor product approximation can be written as

$$\mathcal{A}_N^{\mathbf{i}}[\mathbf{w}_h](\mathbf{y}) = (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})[\mathbf{w}_h](\mathbf{y}) = \sum_{\mathbf{j}\in J} \mathbf{w}_h(\mathbf{y}_{\mathbf{j}})\lambda_{\mathbf{j}}(\mathbf{y}), \tag{4.9}$$

and the approximated solution is given by

$$\mathbf{w}_{h,\eta}(\mathbf{y}) = \mathcal{A}_N^{\mathbf{i}}[\mathbf{w}_h](\mathbf{y}).$$

4.1.3 Sparse grid approximations

The main idea of sparse grid approximations is not to evaluate the function on a single tensor product set as described in Section 4.1.2, but on the union of several smaller tensor product sets which are determined by a rule on the multi-index $\mathbf{i} = (i_1, \ldots, i_N)$. Subsequently, the Smolyak sparse grid will be considered in more detail, in particular the so-called isotropic Smolyak sparse grid. Its construction is given and illustrated with an example. The idea of a quasi optimal sparse grid is shortly pointed out at the end of this subsection.

The Smolyak sparse grid is built via the difference operator

$$\Delta^{i_n} = \mathcal{U}^{i_n} - \mathcal{U}^{i_n - 1},$$

where $\mathcal{U}^{i_n}$ was defined in (4.8), and $\mathcal{U}^0 = 0$.

Let $\ell \in \mathbb{N}$ denote the level. Denote by $\mathbf{I}(\ell) \in \mathbb{N}^N$ the index set corresponding to the level ℓ with

$$\mathbf{I}(\ell) \subset \mathbf{I}(\ell+1), \quad \mathbf{I}(0) = \{\mathbf{1}\} \text{ and } \cup_{\ell \in \mathbb{N}} \mathbf{I}(\ell) = \mathbb{N}^N.$$

The hierarchical surplus operator is defined by

$$\mathbf{\Delta}^{\mathbf{i}} = \bigotimes_{n=1}^{N} (\mathcal{U}^{i_n} - \mathcal{U}^{i_n - 1}).$$

The sparse grid approximation of a function $\mathbf{w}_h \in L^2_\rho(\Gamma; W(D)) \cap C^0_\sigma(\Gamma; W(D))$ at some level $\ell \in \mathbb{N}$ is given by

$$\mathcal{A}^{\mathbf{I}(\ell)} : L^2_\rho(\Gamma; W(D)) \cap C^0_\sigma(\Gamma; W(D)) \to L^2_\rho(\Gamma; W(D)),$$

with

$$\mathcal{A}^{\mathbf{I}(\ell)}[\mathbf{w}_h](\mathbf{y}) = \sum_{\mathbf{i} \in \mathbf{I}(\ell)} \mathbf{\Delta}^{m(\mathbf{i})}[\mathbf{w}_h](\mathbf{y}).$$

The approximated solution is defined as

$$\mathbf{w}_{h,\eta} = \mathcal{A}^{\mathbf{I}(\ell)}[\mathbf{w}_h](\mathbf{y}).$$

In the following, only the isotropic Smolyak sparse grid is considered in detail. Recall that *isotropic* means that every direction n is equally weighted. For a study of the anisotropic sparse grid it is referred to [23] as it is out of the scope of the thesis.

The *isotropic Smolyak approximation* is a sparse grid method, which is used in, for example, [22] and [11].
By the choice

$$\mathbf{I}(\ell) = X(\ell, N) = \left\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1} : \sum_{n=1}^{N} (i_n - 1) \leq \ell \right\}$$

the isotropic Smolyak formula

$$\mathcal{A}^{\mathbf{I}(\ell)} = \mathcal{A}^{X(\ell,N)} = \sum_{\mathbf{i} \in X(\ell,N)} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_N}) \tag{4.10}$$

is obtained. The equivalence of $\mathcal{A}^{X(\ell,N)}$ to

$$\mathcal{A}^{Y(\ell,N)} = \sum_{\mathbf{i} \in Y(\ell,N)} (-1)^{\ell+N-|\mathbf{i}|} \binom{N-1}{\ell+N-|\mathbf{i}|} \cdot (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})$$

is shown in [25, p. 13–14], where

$$Y(\ell, N) = \Big\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1} : \ell - N + 1 \leq \sum_{n=1}^{N} (i_n - 1) \leq \ell \Big\}.$$

Note that $\mathcal{A}^{Y(\ell,N)}$ contains the tensor product $\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N}$ for every index $\mathbf{i} \in Y(\ell, N)$. Because of the equivalence of $\mathcal{A}^{X(\ell,N)}$ and $\mathcal{A}^{Y(\ell,N)}$, for the computation of the isotropic Smolyak formula $\mathcal{A}^{X(\ell,N)}$ it is sufficient to evaluate the function in question on the sparse grid

$$\mathcal{H}(\ell, N) = \bigcup_{\mathbf{i} \in Y(\ell,N)} (\theta^{i_1} \times \cdots \times \theta^{i_N}) \subset \Gamma,$$

which is the union of tensor product sets $(\theta^{i_1} \times \cdots \times \theta^{i_N})$ with indices in $Y(\ell, N)$.

In Figure 1 an example for the construction of the tensor products for $N = 2$, namely $\theta^{i_1} \times \theta^{i_2}$, is given. The abscissae $\theta^{i_1}, \theta^{i_2}$ are built on the Clenshaw-Curtis points with the number of abscissae in each direction $m(i_n)$ determined by (4.6). Each picture in the figure represents $\theta^{i_1} \times \theta^{i_2}$ for the indices (i_1, i_2) given above each picture. For the level $\ell = 3$, the pairs of indices (i_1, i_2) which belong to the set $Y(\ell, N) = Y(3, 2)$ and their tensor grids are highlighted in black.

As the pictures highlighted in black indicate, the grid points belonging to the pairs $(1, 3), (2, 2), (3, 1)$ are again grid points belonging to the pairs $(1, 4)^*$, $(2, 3)^*$, $(3, 2)^*$ due to the nestedness of the Clenshaw-Curtis abscissae. Thus, to build the sparse grid on $Y(\ell, N)$ it is sufficient to consider a smaller set of indices (i_1, i_2) which will be specified in the following.

In general, if the abscissae for the collocation points are nested, the function in question has only to be evaluated on the grid

$$\mathcal{H}(\ell, N) = \bigcup_{\mathbf{i} \in \tilde{X}(\ell,N)} (\theta^{i_1} \times \cdots \times \theta^{i_N}) \subset \Gamma,$$

with

$$\tilde{X}(\ell, N) = \Big\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1} : \sum_{n=1}^{N} (i_n - 1) = \ell \Big\}.$$

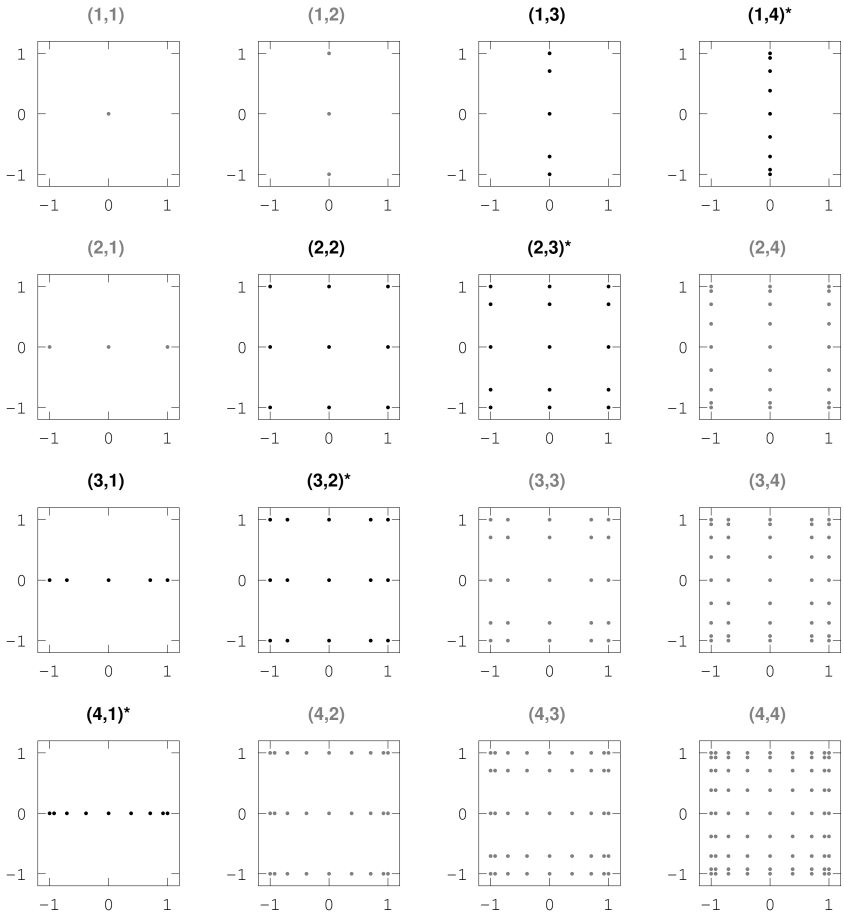

Figure 1: Tensor product grids ($\theta^{i_1} \times \theta^{i_2}$) based on Clenshaw-Curtis points for $N = 2$, indices $i_1, i_2 \in \{1, \dots, 4\}$ and the rule for $m(i_n)$ as in equation (4.6).

In the example given in Figure 1, $\tilde{X}(\ell, N) = \tilde{X}(3, 2)$ is the set of indices $(i_1, i_2)^*$ marked with a *.

The set of collocation points for some level ℓ equals the points of the sparse grid $\mathcal{H}(\ell, N)$. The number of collocation points η of a fixed level ℓ is given by

$$\eta = |\mathcal{H}(\ell, N)|,$$

and the set of collocation points by

$$\{\mathbf{y}_i : \mathbf{y}_i \in \Gamma \subset \mathbb{R}^N\}_{i=1}^{\eta} = \mathcal{H}(\ell, N),$$

where the set $\mathcal{H}(\ell, N)$ has to be chosen according to the case of nested or non-nested abscissae.

In Figure 2, the sparse grid $\mathcal{H}(\ell, N) = \mathcal{H}(3,2)$ is depicted on the left. It is the union of the sets $(\theta^{i_1} \times \theta^{i_2})$ belonging to the pairs of indices $(1,4)^*$, $(2,3)^*$, $(3,2)^*$, and $(4,1)^*$. Note that this grid itself is not again a tensor product grid. By comparing the grid $\mathcal{H}(3,2)$ with the tensor product grid with the pair of indices $(4,4)$ (depicted in the right), which has the same number of maximal grid points in each direction, the reduction of grid points as a result of a sparse grid construction becomes obvious.

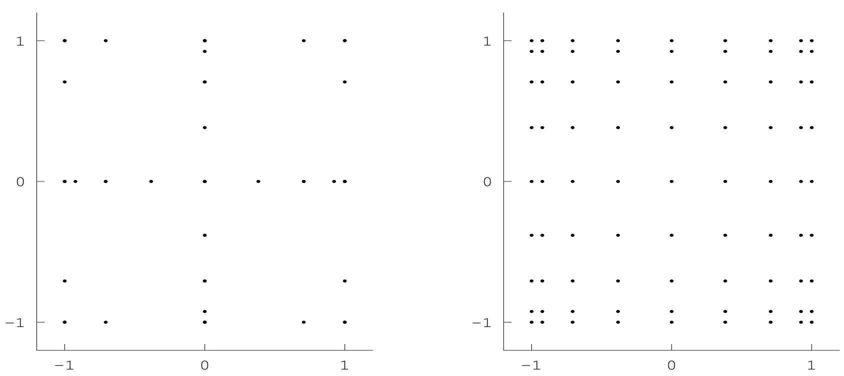

Figure 2: Comparison of sparse grid (left) and tensor product grid (right).

The idea of a *quasi optimal grid* consists in the construction of the index set $\mathbf{I}(\ell)$. The index set is constructed by a greedy algorithm as to minimize the approximation error. Hence, the problem of building a sparse grid is reformulated to a knapsack problem [21, p. 2].

4.2 Numerical integration

Often, it is more of interest to compute the expected value or other statistical quantities like the variance or higher moments of some solution than the solution itself. In these cases, instead of approximating the solution and then taking its expected value, it is possible to directly apply a quadrature rule to determine the expected value.
In this section, in analogue to the approximation of the previous section, tensor product and sparse grid quadrature formulae are introduced. The focus is on the computation of the expected value because formulae for the other moments can be obtained in a similar manner.

Let the function $\mathbf{w}$ be defined on $C^0_\sigma(\Gamma; W(D))$. Its expected value is given by

$$\mathbb{E}[\mathbf{w}](\mathbf{x}) = \int_\Gamma \mathbf{w}(\mathbf{y}, \mathbf{x})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \in W(D).$$

Quadrature formulae are applied to functions $\mathbf{w}$ which should be integrated over some N-dimensional domain Γ, i.e.,

$$I_N\mathbf{w} = \int_\Gamma \mathbf{w}(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

First, consider the one-dimensional case with a univariate function $\mathbf{w}$

$$I_1\mathbf{w} = \int_{\Gamma_n} \mathbf{w}(y)\,\mathrm{d}y$$

with $n \in \{1, \ldots, N\}$.
Since quadrature in the stochastic variable should be applied, let

$$\theta^{i_n} = \{\theta^{i_n}_1, \theta^{i_n}_2, \ldots, \theta^{i_n}_{m(i_n)}\} \subset \Gamma_n$$

be a quadrature point set for each index i_n.
Let

$$Q^{i_n}_1 : C^0_\sigma(\Gamma_n; W(D)) \to W(D)$$

be the quadrature formula corresponding to the index i_n given by

$$Q^{i_n}_1[\mathbf{w}] = \sum_{j_n=1}^{m(i_n)} \omega^{i_n}_{j_n} \mathbf{w}(\theta^{i_n}_{j_n})$$

with the weights $\omega^{i_n}_{j_n}$, which will be specified later. If the point sets $\theta^{i_n} \subset \theta^{i_n+1}$ are nested, one also speaks of a nested quadrature formula.
Two multi-dimensional quadrature formulae shall be given, the tensor product of N quadrature formulae and the Smolyak sparse grid formula. For both, the connection to approximation is pointed out in Section 4.3. Assume that $\mathbf{w}$ is a function depending on N directions. The *tensor product of N quadrature formulae*

$$Q^{\mathbf{i}}_N = (Q^{i_1}_1 \otimes \cdots \otimes Q^{i_N}_1) : C^0_\sigma(\Gamma; W(D)) \to W(D)$$

is defined by

$$Q^{\mathbf{i}}_N[\mathbf{w}] = (Q^{i_1}_1 \otimes \cdots \otimes Q^{i_N}_1)[\mathbf{w}] = \sum_{j_1=1}^{m(i_1)} \cdots \sum_{j_N=1}^{m(i_N)} \omega^{i_1}_{j_1} \cdot \cdots \cdot \omega^{i_N}_{j_N} \mathbf{w}(\theta^{i_1}_{j_1}, \ldots, \theta^{i_N}_{j_N}).$$

Also sparse grid quadrature formulae exist. Define the difference quadrature formula by

$$\Delta_1^{i_n} = Q_1^{i_n} - Q_1^{i_n-1},$$

with $Q_1^0 = 0$. Then, the *Smolyak quadrature formula* for functions $\mathbf{w}$ depending on N dimensions is given by

$$Q_N^{X(\ell,N)} : C_\sigma^0(\Gamma; W(D)) \to W(D),$$

with

$$Q_N^{X(\ell,N)}[\mathbf{w}] = \sum_{\mathbf{i} \in X(\ell,N)} (\Delta_1^{i_1} \otimes \cdots \otimes \Delta_1^{i_N})[\mathbf{w}].$$

Note that here the level may start at 0 and not at 1. This is in contrast to [13], where the general construction of the formulae is taken from and adjusted to the approximations given in the previous subsection.
The quadrature formula $Q_N^{X(\ell,N)}$ is, as before in Section 4.1.3, equivalent to

$$Q_N^{Y(\ell,N)}[\mathbf{w}] = \sum_{\mathbf{i} \in Y(\ell,N)} (-1)^{\ell+N-|\mathbf{i}|} \binom{N-1}{\ell+N-|\mathbf{i}|} \cdot (Q_1^{i_1} \otimes \cdots \otimes Q_1^{i_N})[\mathbf{w}].$$

The two former formulae, i.e., the tensor product and the Smolyak sparse grid quadrature formula, have the same form as the tensor product and Smolyak sparse grid approximation, except that they are formulated for quadrature rules (i.e., $\mathcal{U}^{i_n}$ is replaced by $Q_1^{i_n}$).

4.3 Connection of collocation and quadrature

In the following, the connection of collocation techniques to quadrature is pointed out. For the choice of Gaussian collocation points, both, the tensor product of N quadrature formulae and the Smolyak quadrature, are shown to be equal to the expected value of the solution determined by a tensor product and Smolyak approximation, respectively.
For this, before showing the equivalence for the tensor product and the Smolyak quadrature to the expected value of the respective approximation, the multi-dimensional Gaussian quadrature is introduced.

The multi-dimensional Gaussian quadrature is based on the one-dimensional Gaussian quadrature which is therefore the starting point.
Let the collocation point set (4.5)

$$\theta^{i_n} = \{\theta_1^{i_n}, \theta_2^{i_n}, \ldots, \theta_{m(i_n)}^{i_n}\} \subset \Gamma_n$$

consist of Gaussian points. For example, for the case of the bounded parameter space $\Gamma_n = [-1, 1]$ with uniform density the set can be chosen as the roots of the Legendre polynomials of degree $m(i_n)$, and for the case of the unbounded parameter space $\Gamma_n = (-\infty, \infty)$ with normal density as the roots of the Hermite polynomials of degree $m(i_n)$; see Section 4.1.1 for Gaussian points. Based on these points a Gaussian quadrature can be applied to compute the expectet value. The Gaussian quadrature is known to be exact for all polynomials of degree $2(m(i_n) - 1) + 1 = 2m(i_n) - 1$.

The weights in one dimension are given by

$$\omega_j^{i_n} = \int_{\Gamma_n} \lambda_j^{i_n}(y)\rho_n(y)\,\mathrm{d}y.$$

Recall that

$$J = \{1, \dots, m(i_1)\} \times \dots \times \{1, \dots, m(i_N)\}.$$

The tensor product weights for all $\mathbf{j}$ in J are given by

$$\omega_{\mathbf{j}} = \prod_{k=1}^{N} \omega_{j_k}^{i_k} \quad \text{with} \quad \sum_{\mathbf{j} \in J} \omega_{\mathbf{j}} = 1.$$

For the multi-dimensional integral it follows

$$\omega_{\mathbf{j}} = \int_{\Gamma} \lambda_{\mathbf{j}}(\mathbf{y})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}. \tag{4.11}$$

The property

$$\int_{\Gamma_n} \lambda_k^{i_n}(y)\lambda_j^{i_n}(y)\rho_n(y)\,\mathrm{d}y = \omega_j^{i_n}\delta_{i,j}$$

holds, where $\delta_{i,j}$ is the Kronecker delta, and similarly, in the multi-dimensional case,

$$\int_{\Gamma} \lambda_{\mathbf{j}}(\mathbf{y})\lambda_{\mathbf{j}'}(\mathbf{y})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} = \omega_{\mathbf{j}}\delta_{\mathbf{j},\mathbf{j}'},$$

which means that the Lagrange polynomials are orthogonal to each other. The latter is necessary for a Gaussian quadrature.

Now, the equivalence of collocation and quadrature for the computation of the expected value of the solution with Gaussian points as collocation and quadrature points will be shown. First, consider the tensor product approximation (4.9) by Lagrange polynomials in the parameter direction as in Section 4.1.2, i.e.,

$$\mathbf{w}_{h,\eta}(\mathbf{y}) = \mathcal{A}_N^{\mathbf{i}}[\mathbf{w}_h](\mathbf{y}) = (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})[\mathbf{w}_h](\mathbf{y}) = \sum_{\mathbf{j} \in J} \mathbf{w}_h(\mathbf{y}_{\mathbf{j}}) \lambda_{\mathbf{j}}(\mathbf{y}),$$

with a continuous function $\mathbf{w}_h$ in $C_\sigma^0(\Gamma; \mathbb{R})$.
After having obtained the semidiscrete solution $\mathbf{w}_h$, see (4.2), and the semidiscrete approximations $\mathbf{w}_h(\mathbf{y}_{\mathbf{j}}, \mathbf{x})$, the expected value can be computed. The expected value $\mathbb{E}[\mathbf{w}_{h,\eta}](\mathbf{x})$ can by definition of $\mathbf{w}_{h,\eta}$ and by definition of the weights (4.11) also be obtained by applying a quadrature rule on the semidiscrete approximation:

$$\begin{aligned}
\mathbb{E}[\mathbf{w}_{h,\eta}](\mathbf{x}) &= \mathbb{E}[\mathcal{A}_N^{\mathbf{i}}[\mathbf{w}_h]](\mathbf{x}) = \sum_{\mathbf{j} \in J} \mathbf{w}_h(\mathbf{y}_{\mathbf{j}}, \mathbf{x}) \int_\Gamma \lambda_{\mathbf{j}}(\mathbf{y}) \rho(\mathbf{y})\, \mathrm{d}\mathbf{y} \\
&\overset{(4.11)}{=} \sum_{\mathbf{j} \in J} \mathbf{w}_h(\mathbf{y}_{\mathbf{j}}, \mathbf{x})\, \omega_{\mathbf{j}} = Q_N^{\mathbf{i}}[\mathbf{w}_h](\mathbf{x}),
\end{aligned}$$

which is exactly the tensor product of N quadrature formulae.

In the case of a Smolyak approximation of the semidiscrete solution $\mathbf{w}_h$, i.e.,

$$\begin{aligned}
\mathbf{w}_{h,\eta}(\mathbf{y}, \mathbf{x}) &= \mathcal{A}^{Y(\ell,N)}[\mathbf{w}_h](\mathbf{y}, \mathbf{x}) \\
&= \sum_{\mathbf{i} \in Y(\ell,N)} (-1)^{\ell+N-|\mathbf{i}|} \binom{N-1}{\ell+N-|\mathbf{i}|} \cdot (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})[\mathbf{w}_h](\mathbf{y}, \mathbf{x}),
\end{aligned}$$

the expected value $\mathbb{E}[\mathbf{w}_{h,\eta}](\mathbf{x})$ is given by

$$\begin{aligned}
\mathbb{E}[\mathbf{w}_{h,\eta}](\mathbf{x}) &= \sum_{\mathbf{i} \in Y(\ell,N)} (-1)^{\ell+N-|\mathbf{i}|} \binom{N-1}{\ell+N-|\mathbf{i}|} \cdot \mathbb{E}[(\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_N})\mathbf{w}_h](\mathbf{x}) \\
&= \sum_{\mathbf{i} \in Y(\ell,N)} (-1)^{\ell+N-|\mathbf{i}|} \binom{N-1}{\ell+N-|\mathbf{i}|} \cdot (Q_1^{i_1} \otimes \cdots \otimes Q_1^{i_N})[\mathbf{w}_h](\mathbf{x}) \\
&= Q_N^{Y(\ell,N)}[\mathbf{w}_h](\mathbf{x}).
\end{aligned}$$

The latter is the above defined Smolyak quadrature formula applied to the semidiscrete solution.
Using the result of the tensor product approximation the explicit rule for computing the Smolyak quadrature is obtained:

$$Q_N^{Y(\ell,N)}[\mathbf{w}_h](\mathbf{x}) = \sum_{\mathbf{i} \in Y(\ell,N)} (-1)^{\ell+N-|\mathbf{i}|} \binom{N-1}{\ell+N-|\mathbf{i}|} \sum_{\mathbf{j} \in J} \mathbf{w}_h(\mathbf{y}_{\mathbf{j}}, \mathbf{x}) \omega_{\mathbf{j}}.$$

Remark 4.1 – Variance and higher order moments:
The variance of $\mathbf{w}_{h,\eta}$ is given by

$$\mathbb{V}[\mathbf{w}](\mathbf{x}) = \mathbb{E}[\mathbf{w}^2](\mathbf{x}) - (\mathbb{E}[\mathbf{w}](\mathbf{x}))^2.$$

With the above derived connections of the approximations to quadrature for the variance of $\mathbf{w}_{h,\eta}$ it follows

$$\begin{aligned}\mathbb{V}[\mathbf{w}_{h,\eta}](\mathbf{x}) = \mathbb{V}[\mathcal{A}_N^{\mathbf{i}}[\mathbf{w}_h]](\mathbf{x}) &= \sum_{\mathbf{j}\in J} \mathbf{w}_h^2(\mathbf{y}_\mathbf{j}, \mathbf{x})\,\omega_\mathbf{j} - \left(\sum_{\mathbf{j}\in J} \mathbf{w}_h(\mathbf{y}_\mathbf{j}, \mathbf{x})\,\omega_\mathbf{j}\right)^2 \\ &= Q_N^{\mathbf{i}}[\mathbf{w}_h^2](\mathbf{x}) - \left(Q_N^{\mathbf{i}}[\mathbf{w}_h](\mathbf{x})\right)^2,\end{aligned}$$

and

$$\begin{aligned}\mathbb{V}[\mathbf{w}_{h,\eta}](\mathbf{x}) &= \mathbb{V}[\mathcal{A}^{Y(\ell,N)}[\mathbf{w}_h]](\mathbf{x}) \\ &= Q_N^{Y(\ell,N)}[\mathbf{w}_h^2](\mathbf{x}) - \left(Q_N^{Y(\ell,N)}[\mathbf{w}_h](\mathbf{x})\right)^2.\end{aligned}$$

Similarly, formulae for higher order moments can be derived, see [15, p. 18].□

4.4 Error estimates

This section gives error estimates for the tensor product and for the Smolyak approximation presented before as well as for the tensor product of N quadrature formulae. These estimates are valid for solutions admitting certain analyticity results. They hold in particular for the solution of the diffusion problem and its mixed form.

The error estimates of the approximations are a priori estimates for the total error

$$\mathbf{w} - \mathbf{w}_{h,\eta}$$

in the norm of the space $L^2_\rho(\Gamma; W(D))$ which has to be specified with respect to the Banach space $W(D)$ for the problem that is considered. Note that the results are not derived for the more general space $L^q_\rho(\Gamma; W(D))$, but only for the case $q = 2$, where the space factorizes.
The total error can be split into the discretization error in space and an approximation or interpolation error in the stochastic variable, respectively, i.e.,

$$\mathbf{w} - \mathbf{w}_{h,\eta} = (\mathbf{w} - \mathbf{w}_h) + (\mathbf{w}_h - \mathbf{w}_{h,\eta}).$$

Section 4.3 dealt with the connection to quadrature, and the collocation method could be understood as a quadrature under certain conditions on the collocation and quadrature points. Hence, in the following the results for $\mathbf{w} - \mathbf{w}_{h,\eta}$ will be given using grids in the collocation method that are quadrature points. More precisely, grids with Gaussian abscissae are used. Then, all estimates which are available on the total error, where the approximated solution is obtained by the collocation method, i.e., $\|\mathbf{w} - \mathbf{w}_{h,\eta}\|_{L^2_\rho(\Gamma;W(D))}$, build an upper bound for the quadrature error. This is a result of the following lemma.

Lemma 4.2 – Approximation of expected value:
Let $W(D) = L^2(D)$ or $W(D) = H^1(D)$. Then

$$\|\mathbb{E}[\mathbf{w} - \mathbf{w}_{h,\eta}]\|_{W(D)} \leq \|\mathbf{w} - \mathbf{w}_{h,\eta}\|_{L^2_\rho(\Gamma;W(D))}.$$

□

Proof: The lemma will be proven for $W(D) = L^2(D)$. For $W(D) = H^1(D)$ the statement follows by the same steps.
It is

$$\begin{aligned}
\|\mathbb{E}[\mathbf{w} - \mathbf{w}_{h,\eta}]\|^2_{L^2(D)} &= \left\| \int_\Gamma (\mathbf{w} - \mathbf{w}_{h,\eta})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \right\|^2_{L^2(D)} \\
&= \int_D \left| \int_\Gamma (\mathbf{w} - \mathbf{w}_{h,\eta})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \right|^2 \mathrm{d}\mathbf{x} \\
&\leq \int_D \left(\int_\Gamma |\mathbf{w} - \mathbf{w}_{h,\eta}|\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \right)^2 \mathrm{d}\mathbf{x} \\
&= \int_D \left(\int_\Gamma |\mathbf{w} - \mathbf{w}_{h,\eta}| \cdot 1 \;\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \right)^2 \mathrm{d}\mathbf{x}.
\end{aligned}$$

Denoting $\left(\int_\Gamma |\mathbf{w} - \mathbf{w}_{h,\eta}| \cdot 1 \;\rho(\mathbf{y})\,d\mathbf{y}\right)^2 = \||\mathbf{w} - \mathbf{w}_{h,\eta}| \cdot 1\|^2_{L^1_\rho(\Gamma)}$ and applying the Cauchy-Schwarz inequality (see Lemma B.2) with the functions $|\mathbf{w} - \mathbf{w}_{h,\eta}|$ and 1, it follows

$$\||\mathbf{w} - \mathbf{w}_{h,\eta}| \cdot 1\|^2_{L^1_\rho(\Gamma)} \leq \||\mathbf{w} - \mathbf{w}_{h,\eta}|\|^2_{L^2_\rho(\Gamma)} \;\|1\|^2_{L^2_\rho(\Gamma)}, \tag{4.12}$$

and hence

$$
\begin{aligned}
\|\mathbb{E}[\mathbf{w}-\mathbf{w}_{h,\eta}]\|^2_{L^2(D)} &\leq \int_D \||\mathbf{w}-\mathbf{w}_{h,\eta}|\cdot 1\|^2_{L^1_\rho(\Gamma)}\,\mathrm{d}\mathbf{x} \\
&\overset{(4.12)}{\leq} \int_D \||\mathbf{w}-\mathbf{w}_{h,\eta}|\|^2_{L^2_\rho(\Gamma)}\ \|1\|^2_{L^2_\rho(\Gamma)}\,\mathrm{d}\mathbf{x} \\
&= \int_D \left(\int_\Gamma |\mathbf{w}-\mathbf{w}_{h,\eta}|^2\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\cdot \underbrace{\int_\Gamma 1^2\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}}_{=1}\right)\mathrm{d}\mathbf{x} \\
&= \int_D\int_\Gamma |\mathbf{w}-\mathbf{w}_{h,\eta}|^2\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} \\
&= \int_\Gamma\int_D |\mathbf{w}-\mathbf{w}_{h,\eta}|^2\,\mathrm{d}\mathbf{x}\,\rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= \|\mathbf{w}-\mathbf{w}_{h,\eta}\|^2_{L^2_\rho(\Gamma;L^2(D))},
\end{aligned}
$$

which concludes the proof. The second to the last equality (i.e., the change of the order of integration) holds true by the Theorem of Fubini. ∎

Remark 4.3:
As shown in Section 4.3, the expected value of the solution $\mathbf{w}_{h,\eta}$ equals a Gaussian quadrature rule if the collocation points are chosen to be Gaussian quadrature points, and if the polynomials in the approximation are of Lagrangian type. Hence, by the previous lemma, one gets the following estimate on the Gaussian quadrature, called Q, either being a tensor product or a sparse grid approximation

$$\|\mathbb{E}[\mathbf{w}]-Q[\mathbf{w}_h]\|_{L^2(D)} = \|\mathbb{E}[\mathbf{w}-\mathbf{w}_{h,\eta}]\|_{L^2(D)} \leq \|\mathbf{w}-\mathbf{w}_{h,\eta}\|_{L^2_\rho(\Gamma;L^2(D))}$$

and the same for functions on $H^1_0(D)$.
Moreover, since $H(\mathrm{div};D)\subset[L^2(D)]^d$, the estimates also hold on $H(\mathrm{div};D)$. Hence, the estimates are valid for both, the diffusion problem and its mixed form. □

Although it is of interest to determine the total error $\mathbf{w}-\mathbf{w}_{h,\eta}$ or $\mathbb{E}[\mathbf{w}]-Q[\mathbf{w}_h]$ in the respective norm, the following error estimates are only considered for the approximation or quadrature error in the stochastic variable, i.e., $\mathbf{w}_h-\mathbf{w}_{h,\eta}$ or $\mathbb{E}[\mathbf{w}_h]-Q[\mathbf{w}_h]$. Results on the spatial discretization error follow by finite element theory, see Lemma 5.3 below for the diffusion problem, and Lemma 5.8 for its mixed form. Notice that the derived analyticity results on the solution $\mathbf{w}$ of the diffusion problem and its mixed form also hold on the discretized solution $\mathbf{w}_h$, as it is discretized only in the spatial variable while the stochastic variable stays unchanged (cf. e.g. [24, p. 15]).

4.4.1 Estimate for a tensor product approximation and for a tensor product quadrature

This section gives two estimates, one for a tensor product approximation and the other for a tensor product quadrature, where the abscissae or quadrature points are assumed to be of Gaussian type. Both estimates are valid for bounded or unbounded parameter space Γ, and they are afterwards compared to each other.

The result on the tensor product approximation can be found in [1] and will be stated subsequently. In [1], the approximation quality of the diffusion problem's solution with a uniformly bounded diffusion coefficient is considered. Investigating the steps of the derivation of an estimate it turns out that the result can be generalized to solutions which are functions in $L^2_\rho(\Gamma; W(D))$ as long as they fulfil certain analyticity conditions. The following theorem is thus formulated in the more general manner. Notice that the result holds in particular true for the diffusion problem and its mixed form.
Before stating the theorem on the tensor product approximation, some necessary preliminaries are given. In the tensor product approximation by polynomials the maximal degree of each polynomial in direction $n \in \{1, \dots, N\}$ is contained in the vector $\mathbf{m}(\mathbf{i}) - \mathbf{1} = (m(i_1) - 1, \dots, m(i_N) - 1)$. Assume that all collocation points $\theta^{i_n} \subset \Gamma_n$, $n \in \{1, \dots, N\}$, are of Gaussian type. Recall that these points $\theta_k^{i_n}$ for $k \in \{1, \dots, m(i_n)\}$ are the zeros of polynomials $q \in \mathcal{P}_{m(i_n)}(\Gamma_n)$ which are orthogonal to $\mathcal{P}_{m(i_n)-1}(\Gamma_n)$ with respect to the density ρ_n in $L^2_{\rho_n}(\Gamma_n)$. Further, let for all $y_n \in \Gamma_n$

$$\rho_n(y_n) \leq C_M e^{-(\delta_n y_n)^2}$$

with $C_M > 0$ and

$$\begin{cases} \delta_n > 0 & \text{if } \Gamma_n \text{ is unbounded} \\ \delta_n = 0 & \text{if } \Gamma_n \text{ is bounded} \end{cases},$$

and assume that

$$\sigma_n(y_n) \geq C_m e^{-(\delta_n y_n)^2/4}$$

with $C_m > 0$ (recall that $\sigma_n(y_n) = 1$ for bounded Γ_n).

Theorem 4.4 – Estimate for tensor product approximation:
Assume that the discretized solution $\mathbf{w}_h$ is in $C^0_\sigma(\Gamma; W(D))$ and satisfies the assumptions of Lemma 3.54, i.e., it admits an analytic continuation

$\mathbf{w}_h(y_1,\ldots,y_{n-1},z_n,y_{n+1},\ldots,y_N,\mathbf{x})$, $z_n \in \mathbb{C}$, *in the region of the complex plane*

$$\Sigma(\Gamma_n;\tau_n) = \{z_n \in \mathbb{C} : \mathrm{dist}(z,\Gamma_n) \le \tau_n\}$$

with $0 < \tau_n < \frac{1}{2\gamma_n}$, *and its norm*

$$\|\sigma_n(\operatorname{Re} z_n)\mathbf{w}_h(z_n)\|_{C^0_{\sigma_n^*}(\Gamma_n^*;W(D))} \le C_P(\tau_n,a,\mathbf{f},g)$$

is bounded by a constant $C_P(\tau_n,a,\mathbf{f},g)$.
Then, there exist positive constants r_n, $n \in \{1,\ldots,N\}$, *and a positive constant* C, *which does not depend on the mesh size* h *and the vector containing the approximation polynomial's degree in each direction* $\mathbf{m}(\mathbf{i}) - \mathbf{1}$ *such that the total error can be bounded by*

$$\begin{aligned}\|\mathbf{w}_h - \mathbf{w}_{h,\eta}\|_{L^2_\rho(\Gamma;W(D))} &= \|\mathbf{w}_h - \mathcal{A}^{\mathbf{i}}_N[\mathbf{w}_h]\|_{L^2_\rho(\Gamma;W(D))} \\ &\le C\sum_{n=1}^N \beta_n(m(i_n)-1)e^{-r_n(m(i_n)-1)^{\theta_n}} \max_{z_n\in\Sigma(\Gamma_n;\tau_n)} \sigma_n(\operatorname{Re} z_n)\|\mathbf{w}_h(z_n)\|_{W(D)},\end{aligned}$$

where

$$\begin{cases} \theta_n = \beta_n = 1, \\ r_n = \log\left(\frac{2\tau_n}{|\Gamma_n|}\left(1+\sqrt{1+\frac{|\Gamma_n|^2}{4\tau_n^2}}\right)\right) \end{cases} \quad \text{if } \Gamma_n \text{ is bounded,}$$

and

$$\begin{cases} \theta_n = \frac{1}{2}, \beta_n = \mathcal{O}(\sqrt{m(i_n)-1}), \\ r_n = \tau_n\delta_n \end{cases} \quad \text{if } \Gamma_n \text{ is unbounded.}$$

□

Proof: A proof with uniformly bounded diffusion coefficient a and the Banach space $W(D) = H^1_0(D)$ can be found in [1, pp. 1019–1024]. This proof can be generalized for Banach spaces $W(D)$ by substituting the spaces in the proof in [1] by the Banach space $W(D)$. The formulation on the discretization error given here allows the theorem to hold also for D-bounded diffusion coefficients.
The analyticity assumption on the discrete solution $\mathbf{w}_h$ is required when the best approximation error in each dimension separately is estimated, i.e., for bounded Γ

$$\inf_{\mathbf{v}\in\mathcal{P}_{m(i_n)-1}\otimes W(D)} \|\mathbf{w}_h - \mathbf{v}\|_{C^0(\Gamma_n;W(D))} \le Ce^{-r_n(m(i_n)-1)} \max_{z_n\in\Sigma(\Gamma_n;\tau_n)} \|\mathbf{w}_h(z_n)\|_{W(D)},$$

and for unbounded Γ

$$\inf_{\mathbf{v}\in\mathcal{P}_{m(i_n)-1}\otimes W(D)} \|\mathbf{w}_h - \mathbf{v}\|_{C^0_\sigma(\Gamma_n;W(D))}$$
$$\leq C\mathcal{O}(\sqrt{m(i_n)-1})e^{-r_n\sqrt{m(i_n)-1}} \max_{z_n\in\Sigma(\Gamma_n;\tau_n)} \sigma_n(\operatorname{Re} z_n)\|\mathbf{w}_h(z_n)\|_{W(D)}. \quad \blacksquare$$

It was pointed out in Remark 4.3 that the quadrature error by Gaussian quadrature can be estimated from above by the corresponding estimate of an approximation obtained by the collocation method. Nevertheless it is not advisable to make use of this estimate to, for example, determine the number of quadrature points in order to achieve a certain accuracy. This is because it is not used that Gaussian quadrature (in one direction) is exact for polynomials of degree $2\cdot\tilde{m}-1$, where $\tilde{m}$ denotes the number of quadrature points. The following theorem on an estimate for the tensor product quadrature allows to compare both ways.

Theorem 4.5 – Estimate for tensor product quadrature:
Let the same assumptions as in Theorem 4.4 hold. Let $(\tilde{m}(i_1),\ldots,\tilde{m}(i_N))$ denote a vector containing the number of quadrature points of Gaussian type in directions $n\in\{1,\ldots,N\}$. Then, there exists a constant C independent of $\tilde{m}(i_1)$, $n\in\{1,\ldots,N\}$, and the quadrature error is bounded by

$$\|\mathbb{E}[\mathbf{w}_h] - Q^{\mathbf{i}}_N[\mathbf{w}_h]\|_{W(D)}$$
$$\leq C\sum_{n=1}^{N}\beta_n(2\tilde{m}(i_n)-1)e^{-r_n(2\tilde{m}(i_n)-1)^{\theta_n}} \max_{z_n\in\Sigma(\Gamma_n;\tau_n)} \sigma_n(\operatorname{Re} z_n)\|\mathbf{w}_h(z_n)\|_{W(D)},$$

where

$$\begin{cases} \theta_n = \beta_n = 1, \\ r_n = \log\left(\frac{2\tau_n}{|\Gamma_n|}\left(1+\sqrt{1+\frac{|\Gamma_n|^2}{4\tau_n^2}}\right)\right) \end{cases} \quad \textit{if } \Gamma_n \textit{ is bounded,}$$

and

$$\begin{cases} \theta_n = \frac{1}{2}, \beta_n = \mathcal{O}(\sqrt{2\tilde{m}(i_n)-1}), \\ r_n = \tau_n\delta_n \end{cases} \quad \textit{if } \Gamma_n \textit{ is unbounded.} \quad \square$$

Proof: A proof for Γ_n unbounded is to be found in [15, pp. 12 – 13]. The result for bounded Γ_n is obtained by analogous steps, where only the best approximation error $\min_{\mathbf{v}\in\mathcal{P}_{2\tilde{m}(i_n)-1}\otimes W(D)}\|\mathbf{w}_h - \mathbf{v}\|$ is differently estimated from above (as in the corresponding proof for the tensor product approximation in [1]). $\blacksquare$

Remark 4.6 – Number of quadrature and of collocation points: Let as in [15, p. 12] $\epsilon = \sum_{n=1}^{N} \epsilon_n$ with $\epsilon_n > 0$. Assume that Γ is unbounded, i.e., Γ_n is unbounded for all $n \in \{1, \ldots, N\}$. In order to bound the quadrature error by ϵ, every term of the sum $\sum_{n=1}^{N} \beta_n(2\tilde{m}(i_n)-1)e^{-r_n(2\tilde{m}(i_n)-1)^{\theta_n}}$ has to be bounded by ϵ_n. It is

$$\begin{aligned}\beta_n(2\tilde{m}(i_n)-1)e^{-r_n(2\tilde{m}(i_n)-1)^{\theta_n}} &= \sqrt{2\tilde{m}(i_n)-1}\, e^{-\tau_n\delta_n\sqrt{2\tilde{m}(i_n)-1}} \\ &\leq C_n \cdot e^{-(1-c)\tau_n\delta_n\sqrt{2\tilde{m}(i_n)-1}}\end{aligned}$$

for all $c > 0$ and a constant $C_n > 0$ independent of $\tilde{m}(i_n)$, in particular for $c \in (0,1)$. To achieve an accuracy ϵ_n the following number of quadrature points is required (note that $c \neq 1$):

$$e^{-(1-c)\tau_n\delta_n\sqrt{2\tilde{m}(i_n)-1}} \leq \epsilon_n \Leftrightarrow \tilde{m}(i_n) \geq \frac{1}{2}\left(\frac{(\log\epsilon_n)^2}{(1-c)^2(\tau_n\delta_n)^2} + 1\right).$$

Hence

$$\|\mathbb{E}[\mathbf{w}] - Q_N^{\mathbf{i}}[\mathbf{w}]\|_{W(D)} \leq C \sum_{n=1}^{N} C_n\epsilon_n \leq \tilde{C}\epsilon.$$

If the same accuracy should be gained by applying as upper bound the estimate obtained by the approximation of the collocation method, the sum $\sum_{n=1}^{N} \beta_n(m(i_n)-1)e^{-r_n(m(i_n)-1)^{\theta_n}}$ given in Theorem 4.4 has to be bounded in the same way. Recall that $m(i_n)$ is the number of collocation points in direction n. For every $n \in \{1, \ldots, N\}$ it holds

$$\begin{aligned}\beta_n(m(i_n)-1)e^{-r_n(m(i_n)-1)^{\theta_n}} &= \sqrt{m(i_n)-1}e^{-\tau_n\delta_n\sqrt{m(i_n)-1}} \\ &\leq C_n e^{-(1-c)\tau_n\delta_n\sqrt{m(i_n)-1}}\end{aligned}$$

for all $c > 0$ and a constant $C_n > 0$ independent of $m(i_n)$, again in particular for $c \in (0,1)$. The determination of the number of collocation points to bound the term by ϵ_n gives (note again that $c \neq 1$)

$$e^{-(1-c)\tau_n\delta_n\sqrt{m(i_n)-1}} \leq \epsilon_n \Leftrightarrow m(i_n) \geq \frac{(\log\epsilon_n)^2}{(1-c)^2(\tau_n\delta_n)^2} + 1.$$

Compared to the required number of quadrature points, this number is twice as high.
Hence, to achieve a certain accuracy in estimating the expected value of some solution $\mathbf{w}$ when using collocation points that are Gaussian quadrature points, it is more advisable to use a quadrature formula. □

The remark, indeed, motivates to take quadrature into consideration when computing the expected value of some solution, at least when the abscissae are of Gaussian type.

4.4.2 Estimate for a Smolyak approximation

In this subsection two estimates when using a Smolyak approximation will be given. The first is for a bounded parameter space, the second for an unbounded one. In both cases results on Gaussian abscissae are given.

The following lemma is a modified version of [22, p. 2331, Lemma 3.12]. The modification is necessary because of the reason that has before been mentioned in Remark 3.55. In the proof the term

$$\left\|(I - \mathcal{A}^{X(\ell,N)})[\mathbf{w}_h]\right\|_{L^2_\rho(\Gamma;W(D))}$$

has to be estimated. The dependence on the parameter space is in N dimensions and not only in one dimension. Therefore, the analyticity result on a product subdomain of $\mathbb{C}^N$ is required which was given in Section 3.4.2. The necessary changes in comparison to the proof of Lemma 3.12 in [22, p. 2331] are to be found in Appendix A.

Remark 4.7:
In the following lemma the semidiscrete solution is assumed to factorize, i.e., $\mathbf{w}_h = \mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,N}$. Note that for all functions being finite linear combinations of such functions the lemma can be applied to each summand. The sum of the estimates of each single factorizing function is obtained. $\square$

Lemma 4.8 – Convergence of Smolyak approximation – bdd. Γ:
Let $W(D)$ be a Hilbert space, i.e., a Banach space with a norm induced by a scalar product. Let $\mathbf{w}_h \in C^0(\Gamma; W(D))$ factorize, i.e., $\mathbf{w}_h = \mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,N}$, and admit an extension in the region of the complex plane

$$\Sigma(\Gamma; \boldsymbol{\tau}) = \{\mathbf{z} \in \mathbb{C}^N : \operatorname{dist}(z_n, \Gamma_n) \leq \tau_n, n = 1, \ldots, N\},$$

where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_N) > \mathbf{0}$. Then, the isotropic Smolyak formula (4.10) with Gaussian abscissae chosen according to (4.6) satisfies

$$\|\mathbf{w}_h - \mathcal{A}^{X(\ell,N)}[\mathbf{w}_h]\|_{L^2_\rho(\Gamma;W(D))} \leq \frac{C_1(r)}{2} \frac{1 - C_1(r)^N}{1 - C_1(r)} \max_{\mathbf{z} \in \Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)} \times \begin{cases} e^{-re\log(2)\ell} & \text{if } 0 \leq \ell \leq \frac{N}{\log(2)} \\ e^{-rN2^{\ell/N}} & \text{otherwise.} \end{cases}$$

Hereby, it is

$$C_1(r) = 4 \cdot \frac{8}{e^{2r} - 1} C_2 \left(1 + \frac{1}{\log(2)} \sqrt{\frac{\pi}{2r}}\right),$$

where

$$C_2 = \max_{n=1,\dots,N} \sqrt{\int_{\Gamma_n} \rho_n(y)\, dy} \Bigg(\prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y) \mathrm{d}y} \Bigg)^{1/k}$$

and

$$r = \min_{n=1,\dots,N} \frac{1}{2} \log\Bigg(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4\tau_n^2}{|\Gamma_n|^2}} \Bigg).$$

$\square$

Remark 4.9:
The statement is given for a density function ρ that factorizes. If ρ does not factorize, a factorizing auxiliary density function $\hat{\rho}$ can be introduced. Then,

$$\|\mathbf{w}\|_{L^2_\rho(\Gamma;W(D))} \leq \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma)}^{1/2} \|\mathbf{w}\|_{L^2_{\hat{\rho}}(\Gamma;W(D))} \quad \forall \mathbf{w} \in C^0(\Gamma; W(D)),$$

and the result of the previous lemma has still to be multiplied by $\left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma)}^{1/2}$. $\square$

It is also of interest to have an error bound depending on the number of collocation points $\eta = \eta(\ell, N) = |\mathcal{H}(\ell, N)|$ of the Smolyak sparse grid. Firstly, an increasing number of collocation points involves increasing computational costs. The second reason is a practical issue: It is relevant in the multilevel method which is defined in the next section.

Theorem 4.10 – Convergence w.r.t. $\boldsymbol{\eta}$ – bounded $\boldsymbol{\Gamma}$:
With the assumptions of the previous lemma, Lemma 4.8, for the number of collocation points $\eta = \eta(\ell, N) = |\mathcal{H}(\ell, N)|$ of the Smolyak sparse grid, the following estimates hold

$$\begin{aligned} \|\mathbf{w}_h - \mathcal{A}^{X(\ell,N)}[\mathbf{w}_h]\|_{L^2_\rho(\Gamma;W(D))} &\leq C_1(r) \frac{\max\{1, C_1(r)\}^N}{|1 - C_1(r)|} \max_{\mathbf{z} \in \Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)} \\ &\times \begin{cases} e^{re\log(2)} \eta^{-\mu_1} & \text{if } 0 \leq \ell \leq \frac{N}{\log(2)} \\ e^{-\frac{N\sigma}{2^{1/N}} \eta^{\mu_2}} & \text{otherwise,} \end{cases} \end{aligned}$$

where

$$\mu_1 = \frac{re\log(2)}{\zeta + \log(N)} \quad \text{and} \quad \mu_2 = \frac{\log(2)}{N(\zeta + \log(N))}$$

with $\zeta = 1 + (1 + \log_2(1.5))\log(2)$. $\square$

Remark 4.11 – On the convergence:
Note that for $\ell > \frac{N}{\log(2)}$ the convergence (with respect to η) is subexponential, while for $0 \leq \ell \leq \frac{N}{\log(2)}$ it is only algebraic. But in order to achieve the subexponential convergence for a large number N of dimensions in the parameter space the number of levels has to be quite high. $\square$

The section is concluded by stating a convergence result for an unbounded parameter space Γ. This can be found in [11, p. 11]. The appendix belonging to the proof given there, however, is unclear at some points. Applying the same steps in the proof as given in Appendix A the result of the theorem stays valid (with the additional assumption on $W(D)$ to be a Hilbert space), at least for factorizing functions, and is therefore formulated as this.

Theorem 4.12 – Convergence of Smolyak approx. – Γ unbdd.:
Let $W(D)$ be a Hilbert space, and let $\mathbf{w}_h \in C^0_\sigma(\Gamma; W(D))$ factorize, i.e., $\mathbf{w}_h = \mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,N}$, and admit an analytic extension to the space

$$\Sigma(\Gamma; \boldsymbol{\tau}) = \{\mathbf{z} \in \mathbb{C}^N : \mathrm{dist}(z_n, \Gamma_n) \leq \tau_n, n = 1, \ldots, N\},$$

and let $\mathbf{w}_h \in C^0_\sigma(\Sigma(\Gamma; \boldsymbol{\tau}); W(D))$, i.e.,

$$\max_{\mathbf{z} \in \Sigma(\Gamma;\boldsymbol{\tau})} \sigma(\mathrm{Re}\,\mathbf{z}) \|\mathbf{w}_h(\mathbf{z})\|_{W(D)} < \infty$$

with $\rho(\mathbf{y}) = \prod_{n=1}^N \rho_n(y_n)$ and $\sigma(\mathbf{y}) = \prod_{n=1}^N \sigma_n(y_n)$, where for $n \in \{1, \ldots, N\}$ it is $\rho_n(y_n) = \frac{e^{-y_n^2/2}}{\sqrt{2\pi}} \leq c e^{-\delta_n y_n^2} \sigma_n(y_n)$ with $c > 0$ and $\delta_n > 0$. Then, if the collocation points are chosen according to the rule (4.6), it holds

$$\|\mathbf{w}_h - \mathcal{A}^{X(\ell,N)}[\mathbf{w}_h]\|_{L^2_\rho(\Gamma;W(D))} \leq C(r) \frac{1 - C(r)^N}{1 - C(r)} \max_{\mathbf{z} \in \Sigma(\Gamma;\boldsymbol{\tau})} \sigma(\mathrm{Re}\,\mathbf{z}) \|\mathbf{w}_h(\mathbf{z})\|_{W(D)}$$
$$\times \begin{cases} e^{-\ell \frac{\log(2)}{2} (r \frac{e}{\sqrt{2}} - 1)} & \text{if } 0 \leq \ell \leq \frac{2N}{\log(2)} \\ e^{-r \frac{N}{\sqrt{2}} \sqrt{2^{\ell/N}} + \ell \frac{\log(2)}{2}} & \text{otherwise,} \end{cases}$$

where $C(r)$ is a constant depending on $r = \min_{n=1,\ldots,N} \tau_n$.

In dependence on the number η of collocation points the error can be bounded by

$$\|\mathbf{w}_h - \mathcal{A}^{X(\ell,N)}[\mathbf{w}_h]\|_{L^2_\rho(\Gamma;W(D))} \leq C(r) \frac{1 - C(r)^N}{1 - C(r)} \max_{\mathbf{z} \in \Sigma(\Gamma;\boldsymbol{\tau})} \sigma(\mathrm{Re}\,\mathbf{z}) \|\mathbf{w}_h(\mathbf{z})\|_{W(D)}$$
$$\times \begin{cases} \sqrt{2^{re/\sqrt{2}-1}} \eta^{-\nu_1} & \text{if } 0 \leq \ell \leq \frac{2N}{\log(2)} \\ \frac{\eta^2}{N^2} e^{-\frac{r}{\sqrt{2}} N \eta^{\nu_2}} & \text{otherwise,} \end{cases}$$

where

$$\nu_1 = \frac{\log(2)}{2(\zeta + \log(N))}\left(r\frac{e}{\sqrt{2}} - 1\right) \quad \textit{and} \quad \nu_2 = \frac{\log(2)}{2N(\zeta + \log(N))}$$

with $\zeta = 1 + (1 + \log_2(1.5))\log(2)$. □

Remark 4.13:
The statement given here is slightly modified in comparison to [11, p. 11]. There, it was introduced some variable $R = \sqrt[N]{\tau_1, \ldots, \tau_N}$, and the proof there uses some lemma [22, Lemma A.1]. However, the requirement of the lemma, i.e., $i_m^* - 1 \leq \ell$, where $i_m^* = 1 + \ell/k + \frac{2}{k}\sum_{n=1}^{k} \log_2(\tau_n/\tau_m)$ for $m \in \{1, \ldots, k\}$, $k \in \{1, \ldots, N\}$, is not necessarily satisfied: $\log_2(\tau_n/\tau_m)$ might be for $\tau_n \gg \tau_m$ and small ℓ such that $\ell/k + \frac{2}{k}\sum_{n=1}^{k} \log_2(\tau_n/\tau_m) > \ell$.
By introducing $r = \min_{1,\ldots,N} \tau_n$ the proof is almost the same as the one given in their earlier work [10], and the result is obtained by following the steps of the proof in [10, pp. 15 – 17]. □

Remark 4.14 – Advantage of Smolyak approximation:
A tensor product approximation (or quadrature) suffers from the so-called curse of dimensionality. This expression describes the drastic rise of computational effort (measured in the number of function evaluations) when increasing the number of dimensions. As remarked in [22, p. 2329], the tensor product approximation (and hence also the tensor product quadrature) have a convergence rate bounded by $C(r;N)\eta^{-\frac{r}{N}}$. By increasing the number of dimensions N the rate of convergence decreases. This phenomenon is diminished for a Smolyak approximation. Its convergence rate is bounded by $C(r;N)\eta^{-\frac{r}{\log N}}$, and hence the decrease of the convergence rate is slowed down. For the same accuracy the Smolyak approximation is computationally less costly than a tensor product approximation.
The same should hold for quadrature. Its error estimates for the tensor product grid are derived in a similar manner as for the tensor product approximation, and presumably also the Smolyak sparse quadrature error can be derived similar to the Smolyak approximation. □

5 Multilevel approximation and quadrature

In this section another kind of sparse grid will be described. While the former sparse grids were constructed only on the parameter space Γ, the discretization of the Banach space $W(D)$ was assumed to be given and was not considered furthermore.
The idea of multilevel approximation respectively quadrature is to combine the discretization of the spatial and of the stochastic variable in such a way: Whenever the spatial variable is evaluated on a coarse grid, for the stochastic variable an approximation operator respectively quadrature rule with a high accuracy is applied, and vice versa.

5.1 Discretization of the spatial domain

For both, the multilevel approximation and quadrature the spatial domain has to be discretized. In Section 4 only one nested subspace $W_h(D) \subset W(D)$ was considered. Now, a family $\{W_{h_k}(D)\}_{k\geq 0} \subset W_h(D)$ of finite element spaces with

$$W_{h_k}(D) \subset W_{h_{k+1}}(D), \quad k = 0, 1, 2, \ldots$$

is introduced. Define

$$W_{h_k}(D) = \{\mathbf{v} \text{ continuous on } D : \mathbf{v}|_{\partial D} = \mathbf{0} \text{ and } \mathbf{v}|_\tau \in \mathcal{P}_p \text{ for all } \tau \in \mathcal{T}_{h_k}\},$$

where $\mathcal{P}_p$ consists of the polynomials of some total degree p with $p \geq 1$ and $\mathcal{T}_{h_k}$ for $k \geq 0$ is a triangulation of the domain D with the initial coarse grid triangulation $\mathcal{T}_{h_0}$ which is uniformly refined. For example, a refinement is given with a diameter $h_k = 2^{-k}h_0$ of each simplex in $\mathcal{T}_{h_k}$.
Let for each discretization level h_k and $\mathbf{y} \in \Gamma$

$$\begin{aligned} G_{h_k}(\mathbf{y}) : W(D) &\to W_{h_k}(D) \\ \mathbf{v} &\mapsto \mathbf{v}_{h_k} \end{aligned}$$

be the Galerkin projection of $\mathbf{v}(\mathbf{y}, \cdot)$ to the space $W_{h_k}(D)$ yielding the discretization $\mathbf{v}_{h_k}(\mathbf{y}, \cdot)$ with respect to the spatial variable.

5.2 Approximation in the stochastic variable

For the approximation in the stochastic variable two possibilities are considered. Firstly, the approximation by interpolation operators, and secondly by quadrature rules.

If the interest is in the solution itself, the approximation by interpolation operators should be chosen. If the expected value of the solution has to be approximated, a quadrature rule can be directly applied.
Here, interpolation operators or quadrature formulae depending on the number of grid points are considered. Let

$$\{\mathcal{A}_{\eta_k}\}_{k\geq 0}$$

be a sequence of interpolation operators

$$\mathcal{A}_{\eta_k} : C^0_\sigma(\Gamma; W(D)) \to L^2_\rho(\Gamma; W(D))$$

using η_k grid points. Possible choices for operators are a tensor product or a sparse grid approximation operator like a Smolyak approximation operator (see Section 4.1).
For the choice of a quadrature formula, let

$$\{Q_{\eta_k}\}_{k\geq 0}$$

be a sequence of quadrature formulae

$$Q_{\eta_k} : C^0_\sigma(\Gamma; W(D)) \to W(D)$$

using η_k grid points. The quadrature formulae can be chosen as tensor product quadrature formulae or sparse grid formulae (see Section 4.2).

5.3 Multilevel method

The multilevel method can be understood as a sparse grid method, where the spatial discretization and the approximation in the stochastic variable are composed such that a sparse grid is constructed.
The multilevel approximation of a function $\mathbf{v} \in C_\sigma(\Gamma; W(D))$ for $K \geq 0$ is given by

$$\begin{aligned} \mathbf{v}_K^{ML} &= \sum_{k=0}^{K} \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}] \\ &= \sum_{k=0}^{K} \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k}] - \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_{k-1}}], \end{aligned} \tag{5.1}$$

where $\mathbf{v}_{h_{-1}} = \mathbf{0}$ and the $\mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k}]$ are the single-level approximations discussed in Section 4.1. The variable h_K indicates the finest grid and η_K the maximal number of grid points. The superscript ^{ML} denotes the multilevel approximation.

The definition of the multilevel method shows that to the solution on the coarsest grid with respect to the spatial variable (i.e., $\mathbf{v}_{h_0}$) the most accurate approximation in the stochastic variable (namely $\mathcal{A}_{\eta_K}$) is applied and vice versa. That is why the multilevel method can be understood as a sparse grid approximation. This is in contrast to a full approximation, where the most accurate approximation in the stochastic variable is simply applied to a fine grid with respect to the spatial variable, i.e., $\mathcal{A}_{\eta_K}[\mathbf{v}_{h_K}]$.

The multilevel quadrature of a function $\mathbf{v} \in C_\sigma(\Gamma; W(D))$ for $K \geq 0$ is given by

$$\begin{aligned} \mathbf{v}_K^{MLQ} &= \sum_{k=0}^{K} Q_{\eta_{K-k}}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}] \\ &= \sum_{k=0}^{K} Q_{\eta_{K-k}}[\mathbf{v}_{h_k}] - Q_{\eta_{K-k}}[\mathbf{v}_{h_{k-1}}], \end{aligned} \tag{5.2}$$

where $Q_{\eta_{K-k}}[\mathbf{v}_{h_k}]$ are the single-level quadrature formulae given in Section 4.2 and again $\mathbf{v}_{h_{-1}} = \mathbf{0}$. The superscript ^{MLQ} denotes the multilevel quadrature. It is easy to see that approximation and quadrature are defined analogously, as in the case of the single-level methods.

Again, the connection of these methods consists when choosing the points of the approximations $\mathcal{A}_{\eta_K}$ as quadrature points. If then the expected value of the approximated solution $\mathbb{E}[\mathbf{v}_K^{ML}]$ is considered, by the results of Section 4.3 it follows

$$\begin{aligned} \mathbb{E}[\mathbf{v}_K^{ML}] &= \mathbb{E}\Bigg[\sum_{k=0}^{K} \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k}] - \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_{k-1}}]\Bigg] \\ &= \sum_{k=0}^{K} \mathbb{E}[\mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k}]] - \mathbb{E}[\mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_{k-1}}]] \\ &= \sum_{k=0}^{K} Q_{\eta_{K-k}}[\mathbf{v}_{h_k}] - Q_{\eta_{K-k}}[\mathbf{v}_{h_{k-1}}] \\ &= \mathbf{v}_K^{MLQ}, \end{aligned}$$

where of course the quadrature formula is chosen adequately to the approximation operator (i.e., tensor product quadrature if the approximation operator is a tensor product one, etc.).

5.4 Error estimates

With the $H^2(D)$-regularity at hand, the following result on the finite element solution u_{h_k} of the diffusion problem (3.5) with uniformly bounded diffusion coefficient is given in [16, p. 7].

Lemma 5.1 – Galerkin approximation error – a uniformly bdd.:
Let the domain D be polygonal or polyhedral and convex, and let $f \in L^2(D)$. Then the unique finite element solution u_{h_k} of the diffusion problem (3.5) admits the following a-priori error bound

$$\|u(\mathbf{y}) - u_{h_k}(\mathbf{y})\|_{H^1(D)} \leq C_1 h_k \|u(\mathbf{y})\|_{H^2(D)}, \tag{5.3}$$

where the constant $C_1 > 0$ depends on a_{min} and a_{max}, but not on $\mathbf{y} \in \Gamma$. □

Remark 5.2 – Result for D-bounded a:
If the diffusion coefficient is D-bounded, the same error estimate as in the previous lemma holds with the constant C_1 depending on $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$, and hence on the stochastic variable $\mathbf{y} \in \Gamma$. Thus $C_1 = C_1(\mathbf{y})$. In [15, p. 16] the constant $C_1(\mathbf{y})$ is even specified. It is given by $C_1(\mathbf{y}) = C\sqrt{\frac{a_{max}(\mathbf{y})}{a_{min}(\mathbf{y})}}$ with a constant $C > 0$ independent of $\mathbf{y}$, $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$. □

Lemma 5.3 – $L^r_\rho(\Gamma; H^1_0(D))$ approximation error:
Let the assumptions of the previous lemma hold as well as the assumptions of Theorem 3.64, i.e., it is $u \in L^r_\rho(\Gamma; H^1_0 \cap H^2)$. Then,

$$\|u - u_{h_k}\|_{L^r_\rho(\Gamma; H^1_0(D))} \leq C_1 h_k \|u\|_{L^r_\rho(\Gamma; H^2(D))}.$$

Further, for the D-bounded diffusion coefficient, let still the assumptions of the previous lemma and the assumptions of Theorem 3.66 hold such that the solution $u \in L^{rs}_\rho(\Gamma; H^1_0 \cap H^2)$ for some $1 < s < \infty$, $2 \leq r < \infty$ and let $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$ be regular enough. Then,

$$\|u - u_{h_k}\|_{L^r_\rho(\Gamma; H^1_0(D))} \leq C h_k \|u\|_{L^{rs}_\rho(\Gamma; H^2(D))}.$$

□

Proof: For the uniformly bounded diffusion coefficient it is by equation (5.3)

$$\begin{aligned}
\|u - u_{h_k}\|^r_{L^r_\rho(\Gamma; H^1_0(D))} &= \int_\Gamma \|u(\mathbf{y}) - u_{h_k}(\mathbf{y})\|^r_{H^1_0(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{(5.3)}{\leq} \int_\Gamma C_1^r h_k^r \|u(\mathbf{y})\|^r_{H^2(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&= (C_1 h_k)^r \|u\|^r_{L^r_\rho(\Gamma; H^2(D))}.
\end{aligned}$$

This gives the first estimate.
The statement for the D-bounded diffusion coefficient can be similarly derived applying a Hölder inequality (see Lemma B.2) with $\frac{1}{s} + \frac{1}{q} = 1$.

$$\begin{aligned}
\|u - u_{h_k}\|^r_{L^r_\rho(\Gamma;H^1_0(D))} &= \int_\Gamma \|u(\mathbf{y}) - u_{h_k}(\mathbf{y})\|^r_{H^1_0(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{(5.3)}{\leq} \int_\Gamma C_1^r(\mathbf{y}) h_k^r \|u(\mathbf{y})\|^r_{H^2(D)} \rho(\mathbf{y})\,\mathrm{d}\mathbf{y} \\
&\overset{\text{Lemma B.2}}{\leq} h_k^r \left(\int_\Gamma C_1^{rq}(\mathbf{y})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\right)^{1/q} \left(\int_\Gamma \|u(\mathbf{y})\|^{rs}_{H^2(D)}\right)^{1/s} \\
&\leq h_k^r C^r \|u\|^r_{L^{rs}_\rho(\Gamma;H^2(D))},
\end{aligned}$$

where the constant $C < \infty$ is obtained by the assumption on the data $a_{min}(\mathbf{y})$, $a_{max}(\mathbf{y})$ to be sufficiently regular. ■

Assumption 5.4 – Galerkin approximation error estimates:
Assume that there is an a-priori error estimate of the unique finite element solution $(p_{h_k}, \mathbf{u}_{h_k})$ of the diffusion problem's mixed form (3.9) in the form

$$\|\mathbf{u}(\mathbf{y}) - \mathbf{u}_{h_k}(\mathbf{y})\|_{L^2(D)} \leq c h_k \|\mathbf{u}(\mathbf{y})\|_{H^1(D)}, \tag{5.4}$$
$$\|\nabla \cdot (\mathbf{u}(\mathbf{y}) - \mathbf{u}_{h_k}(\mathbf{y}))\|_{L^2(D)} \leq c h_k |\nabla \cdot \mathbf{u}(\mathbf{y})|_{H^1(D)}, \tag{5.5}$$
$$\|p(\mathbf{y}) - p_{h_k}(\mathbf{y})\|_{L^2(D)} \leq c h_k \big(\|\mathbf{u}(\mathbf{y})\|_{H^1(D)} + \|p(\mathbf{y})\|_{H^1(D)}\big), \tag{5.6}$$

where the constant $c > 0$ depends on a_{min} and a_{max}, and in the case of a D-bounded diffusion coefficient on $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$. The semi-norm $|v|_{H^1(D)}$ is defined as $|v|^2_{H^1(D)} = \sum_{|\alpha|=1} |D^\alpha v|^2_{L^2(D)}$. □

Remark 5.5:
The estimates given in the previous assumption are sensible. Examples for finite element spaces satisfying the above required estimates are to be found in [6, p. 132, Proposition 3.9, and p. 139, Proposition 1.2]. □

Definition 5.6 – The space $H^1(\mathrm{div}; D)$:
Define

$$H^1(\mathrm{div}; D) = \{\mathbf{v} \in [H^1(D)]^d : \nabla \cdot \mathbf{v} \in H^1(D)\}$$

equipped with the norm

$$\|\mathbf{v}\|_{H^1(\mathrm{div};D)} = \sqrt{\|\mathbf{v}\|^2_{H^1(D)} + |\nabla \cdot \mathbf{v}|^2_{H^1(D)}}.$$ □

Remark 5.7:
Actually, an estimate on $\mathbf{u}(\mathbf{y})$ in the norm $\|\cdot\|_{H(\mathrm{div};D)}$ is needed. This can be derived from the estimates in the previous assumption along with the definition of the $H(\mathrm{div}; D)$-norm by

$$\begin{aligned}
&\|\mathbf{u}(\mathbf{y}) - \mathbf{u}_{h_k}(\mathbf{y})\|_{H(\mathrm{div};D)} \\
&\qquad = \sqrt{\|\mathbf{u}(\mathbf{y}) - \mathbf{u}_{h_k}(\mathbf{y})\|^2_{L^2(D)} + \|\nabla \cdot (\mathbf{u}(\mathbf{y}) - \mathbf{u}_{h_k}(\mathbf{y}))\|^2_{L^2(D)}} \\
&\qquad \overset{(5.4),\ (5.5)}{\leq} \sqrt{c^2 h_k^2 \|\mathbf{u}(\mathbf{y})\|^2_{H^1(D)} + c^2 h_k^2 |\nabla \cdot \mathbf{u}(\mathbf{y})|^2_{H^1(D)}} \\
&\qquad = c h_k \|\mathbf{u}(\mathbf{y})\|_{H^1(\mathrm{div};D)}.
\end{aligned}$$

□

Lemma 5.8 – $L^r_\rho(\Gamma; L^2(D) \times H(\mathrm{div}; D))$ approximation error:
Let Assumption 5.4 hold, and let $\mathbf{w} = (p, \mathbf{u}) \in L^2_\rho(\Gamma; H^1(D) \times H^1(\mathrm{div}; D))$. *Then for the mixed form with uniformly bounded diffusion coefficient it is*

$$\|\mathbf{w} - \mathbf{w}_{h_k}\|_{L^2_\rho(\Gamma;L^2(D)\times H(\mathrm{div};D))} \leq C h_k \|\mathbf{w}\|_{L^2_\rho(\Gamma;H^1(D)\times H^1(\mathrm{div};D))}.$$

For the D-bounded diffusion coefficient let also Assumption 5.4 hold. Let $\mathbf{w} = (p, \mathbf{u}) \in L^s_\rho(\Gamma; H^1(D)\times H^1(\mathrm{div}; D))$ *for some* $s \in (2, \infty)$, *and let* $a_{min}(\mathbf{y})$ *and* $a_{max}(\mathbf{y})$ *be regular enough. Then,*

$$\|\mathbf{w} - \mathbf{w}_{h_k}\|_{L^2_\rho(\Gamma;L^2(D)\times H(\mathrm{div};D))} \leq c h_k \|\mathbf{w}\|_{L^s_\rho(\Gamma;H^1(D)\times H^1(\mathrm{div};D))}.$$

□

Proof: The proof is similar to the proof of Lemma 5.3 for the diffusion problem. Therefore, it is omitted here. Nevertheless, it shall be noted that the estimates can be shown to hold for p and $\mathbf{u}$ on their own, where in the proof for p equation (3.22) on page 46 by may be used. ■

Remark 5.9:
With the assumptions of Theorems 3.67 and 3.70 the statements of the previous lemma on p is fulfilled. □

After these preparations the convergence of the multilevel method can be analysed. This will be done both for the collocation method and quadrature in the context of the diffusion problem and its mixed form.

The a-priori error of the multilevel approximation $\mathbf{v}_K^{ML}$ of the solution $\mathbf{v}$ can be split into the spatial discretization error and the stochastic approximation error. The latter is measured in the natural norm $\|\cdot\|_{L^2_\rho(\Gamma;W(D))}$, i.e.,

$$\|\mathbf{v} - \mathbf{v}_K^{ML}\|_{L^2_\rho(\Gamma;W(D))} \leq \|\mathbf{v} - \mathbf{v}_{h_K}\|_{L^2_\rho(\Gamma;W(D))} + \|\mathbf{v}_{h_K} - \mathbf{v}_K^{ML}\|_{L^2_\rho(\Gamma;W(D))}.$$

The first term can be estimated by applying results from finite element theory. The second term, using the telescopic sum (recall that $\mathbf{v}_{h_{-1}} = 0$)

$$\mathbf{v}_{h_K} = \sum_{k=0}^{K} (\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}})$$

and recalling the definition of $\mathbf{v}_K^{ML}$ in equation (5.1), can be estimated by the triangle inequality as follows (see [24, p. 7])

$$\begin{aligned}\|\mathbf{v}_{h_K} - \mathbf{v}_K^{ML}\|_{L^2_\rho(\Gamma;W(D))} &= \left\| \sum_{k=0}^{K} (\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}) - \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}] \right\|_{L^2_\rho(\Gamma;W(D))} \\ &\leq \sum_{k=0}^{K} \|(\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}) - \mathcal{A}_{\eta_{K-k}}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}]\|_{L^2_\rho(\Gamma;W(D))}.\end{aligned}$$

Each term of the sum has to be estimated by the error of a single level approximation of $\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}$, where for example the error bounds for a tensor product or Smolyak approximation given in Section 4.4 can be used. Further analysis on this for the diffusion problem will be carried out below.

In the same way the a-priori error of the multilevel quadrature $\mathbf{v}_k^{MLQ}$ of the solution $\mathbf{v}$ can be split into the spatial discretization error and the stochastic quadrature error. The latter is measured in the natural norm $\|\cdot\|_{W(D)}$, i.e.,

$$\|\mathbb{E}[\mathbf{v}] - \mathbf{v}_K^{MLQ}\|_{W(D)} \leq \|\mathbb{E}[\mathbf{v}] - \mathbb{E}[\mathbf{v}_{h_K}]\|_{W(D)} + \|\mathbb{E}[\mathbf{v}_{h_K}] - \mathbf{v}_K^{MLQ}\|_{W(D)}.$$

The first term, if $W(D) = L^2(D)$ or $W(D) = H^1(D)$, can be approximated by

$$\|\mathbb{E}[\mathbf{v}] - \mathbb{E}[\mathbf{v}_{h_K}]\|_{W(D)} \leq \|\mathbf{v} - \mathbf{v}_{h_K}\|_{L^2_\rho(\Gamma;W(D))}$$

according to Lemma 4.2.
The second term, using again the telescopic sum and the definition of the multilevel quadrature given in equation (5.2), is estimated by

$$\begin{aligned}\|\mathbb{E}[\mathbf{v}_{h_K}] - \mathbf{v}_K^{MLQ}\|_{W(D)} &= \left\| \sum_{k=0}^{K} \mathbb{E}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}] - Q_{\eta_{K-k}}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}] \right\|_{W(D)} \\ &\leq \sum_{k=0}^{K} \|\mathbb{E}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}] - Q_{\eta_{K-k}}[\mathbf{v}_{h_k} - \mathbf{v}_{h_{k-1}}]\|_{W(D)},\end{aligned}$$

where as in the multilevel approximation the terms of the sum can be estimated by using single level results.

For the diffusion problem and its mixed form the following theorem can be concluded by the above given estimates.

Theorem 5.10 – Multilevel approximation and quadrature:
Let $\mathbf{w} \in C^0_\sigma(\Gamma; W(D))$ be the unique weak solution to the diffusion problem (3.5) or its mixed form (3.9), i.e., $W(D) = H^1_0(D)$ or $W(D) = H(\mathrm{div}; D)$. Let the assumptions of Lemma 5.3 and Lemma 5.8 hold true for $r = 2$. For the multilevel approximation, assume that the interpolation operators $\mathcal{A}_{\eta_{K-k}}$ and the number of collocation points are chosen such that

$$\sum_{k=0}^{K} \|(\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}) - \mathcal{A}_{\eta_{K-k}}[\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}]\|_{L^2_\rho(\Gamma;W(D))} \leq ch_K,$$

where $c > 0$ is a constant independent of the mesh widths $h_k, k \in \{0, \dots, K\}$. Then, the multilevel approximation error is bounded by

$$\|\mathbf{w} - \mathbf{w}^{ML}_{h_K}\|_{L^2_\rho(\Gamma;W(D))} \leq Ch_K$$

with a constant $C > 0$ independent of $h_k, k \in \{0, \dots, K\}$.
For the multilevel quadrature, assume that the quadrature formulae $Q_{\eta_{K-k}}$ and the number of quadrature points are chosen such that

$$\sum_{k=0}^{K} \|\mathbb{E}[\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}] - Q_{\eta_{K-k}}[\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}]\|_{W(D)} \leq ch_K$$

with a constant $c > 0$ independent of $h_k, k \in \{0, \dots, K\}$.
The multilevel quadrature error is bounded by

$$\|\mathbb{E}[\mathbf{w}] - \mathbf{w}^{MLQ}_{h_K}\|_{W(D)} \leq Ch_K$$

with another constant $C > 0$ independent of $h_k, k \in \{0, \dots, K\}$. □

Remark 5.11:
The previous theorem arises the question which interpolation operators and which quadrature formulae can be chosen to fulfil the assumptions

$$\sum_{k=0}^{K} \|(\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}) - \mathcal{A}_{\eta_{K-k}}[\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}]\|_{L^2_\rho(\Gamma;W(D))} \leq ch_K$$

and

$$\sum_{k=0}^{K} \|\mathbb{E}[\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}] - Q_{\eta_{K-k}}[\mathbf{w}_{h_k} - \mathbf{w}_{h_{k-1}}]\|_{W(D)} \leq ch_K,$$

respectively. The inequalities have to be verified for the specific choice of an operator or quadrature formula. □

Example 5.12 – Approximation with a Smolyak grid, cf. [24]:
An example that fulfils the assumptions of the previous theorem will be given. Hereby, for the approximation in the stochastic variable a Smolyak sparse grid is chosen. The example closely follows the procedure in [24]. The diffusion coefficient is assumed to be uniformly bounded, and only the diffusion problem in its standard form is considered.
The assumption which essentially has to be verified is

$$\sum_{k=0}^{K} \|(u_{h_k} - u_{h_{k-1}}) - \mathcal{A}_{\eta_{K-k}}[u_{h_k} - u_{h_{k-1}}]\|_{L^2_\rho(\Gamma;H^1_0(D))} \leq ch_K,$$

where $\mathcal{A}_{\eta_{K-k}}$ is chosen as the Smyolak formula $\mathcal{A}^{X(\ell,N)}$ with η_{K-k} collocation points which are chosen according to (4.6) given in Section 4.1.3.
In Theorem 4.10 for bounded and in Theorem 4.12 for unbounded Γ the convergence of the Smolyak approximation with respect to the collocation points has been shown to be at least algebraically, i.e.,

$$\begin{aligned}
&\|(u_{h_k} - u_{h_{k-1}}) - \mathcal{A}_{\eta_{K-k}}[u_{h_k} - u_{h_{k-1}}]\|_{L^2_\rho(\Gamma;H^1_0(D))} \\
&\leq C_1 \eta_{K-k}^{-\mu} \max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \sigma(\operatorname{Re}\mathbf{z}) \|u_{h_k}(\mathbf{z}) - u_{h_{k-1}}(\mathbf{z})\|_{H^1_0(D)}.
\end{aligned}$$

Extending estimate (5.3) to variables $\mathbf{z} \in \Sigma(\Gamma;\boldsymbol{\tau})$ and using the triangle inequality

$$\|u_{h_k}(\mathbf{z}) - u_{h_{k-1}}(\mathbf{z})\|_{H^1_0(D)} \leq \|u(\mathbf{z}) - u_{h_k}(\mathbf{z})\|_{H^1_0(D)} + \|u(\mathbf{z}) - u_{h_{k-1}}(\mathbf{z})\|_{H^1_0(D)}$$

the following estimate is gained:

$$\max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \sigma(\operatorname{Re}\mathbf{z}) \|u_{h_k}(\mathbf{z}) - u_{h_{k-1}}(\mathbf{z})\|_{H^1_0(D)} \leq C_2 h_k \max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \sigma(\operatorname{Re}\mathbf{z}) \|u(\mathbf{z})\|_{H^2(D)}.$$

In this estimate the independence on $\mathbf{z}$ of the constant $C_2 > 0$ and the mesh size h_k is used. If a D-bounded diffusion coefficient is assumed, the constant C_2 depends on $\mathbf{z}$. Thus, not only $\sigma(\operatorname{Re}\mathbf{z})\|u(\mathbf{z})\|_{H^2(D)}$ is maximized, but $C_2(\mathbf{z})\sigma(\operatorname{Re}\mathbf{z})\|u(\mathbf{z})\|_{H^2(D)}$. This case is not considered here, but it is referred to the following Remark 5.13 how this could be handled.

Hence,

$$\begin{aligned}&\|(u_{h_k}-u_{h_{k-1}})-\mathcal{A}_{\eta_{K-k}}[u_{h_k}-u_{h_{k-1}}]\|_{L^2_\rho(\Gamma;H^1_0(D))}\\&\le C_1\eta_{K-k}^{-\mu}C_2h_k\max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})}\sigma(\operatorname{Re}\mathbf{z})\|u(\mathbf{z})\|_{H^2(D)}.\end{aligned}$$

If the solution u is in $C^0_\sigma(\Sigma(\Gamma;\boldsymbol{\tau});H^2(D))$ – this can be derived stating similar assumptions as in Assumption 3.24 and using estimate (3.19) in the proof of Lemma 3.64 - the term

$$\max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})}\sigma(\operatorname{Re}\mathbf{z})\|u(\mathbf{z})\|_{H^2(D)}$$

can be bounded by another constant $C_3>0$ independent of h_k, $k\in\{1,\dots,K\}$. Thus,

$$\begin{aligned}&\|(u_{h_k}-u_{h_{k-1}})-\mathcal{A}_{\eta_{K-k}}[u_{h_k}-u_{h_{k-1}}]\|_{L^2_\rho(\Gamma;H^1_0(D))}\\&\le C_1\eta_{K-k}^{-\mu}C_2h_kC_3.\end{aligned}$$

By choosing

$$\eta_{K-k}^{-\mu}\le\frac{1}{(K+1)C_1C_2C_3}\frac{h_K}{h_k}$$

it is

$$\begin{aligned}&\sum_{k=0}^K\|(u_{h_k}-u_{h_{k-1}})-\mathcal{A}_{\eta_{K-k}}[u_{h_k}-u_{h_{k-1}}]\|_{L^2_\rho(\Gamma;H^1_0(D))}\\&\qquad\le\sum_{k=0}^K h_kC_1C_2C_3\frac{1}{(K+1)C_1C_2C_3}\frac{h_K}{h_k}=h_K.\end{aligned}$$

As by Lemma 5.3 it is $\|u-u_{h_K}\|_{L^2_\rho(\Gamma;H^1_0(D))}\le C_4h_K$, the estimate

$$\|u-u^{ML}_{h_K}\|_{L^2_\rho(\Gamma;H^1_0(D))}\le C_4h_K+h_K=Ch_K$$

is obtained. □

Remark 5.13 – Norm definition:
Probably, it is possible to shift the dependence of the constant on $a_{max}(\mathbf{y})$ and $a_{min}(\mathbf{y})$ to an appropriate norm similar to [20]. Estimates with a constant independent of $a_{max}(\mathbf{y})$ and $a_{min}(\mathbf{y})$ are derived in [20] for the diffusion problem with a diffusion coefficient which is bounded by random variables. The norm which is used there is

$$\|u\|_{L^2_{\tilde{a}\rho}(\Gamma;H^1_0(D))}=\left(\int_\Gamma\|u\|^2_{H^1_0(D)}\tilde{a}(\mathbf{y})\rho(\mathbf{y})\,\mathrm{d}\mathbf{y}\right)^{1/2},$$

where $\tilde{a}$ depends on $a_{min}(\mathbf{y})$ and $a_{max}(\mathbf{y})$. □

5.5 Computational costs

Computational costs are worthwhile considering in the context of the multilevel method to compare the performance of multilevel and single-level methods. Subsequently, it will be shown for the approximation by the collocation method that the computational cost can be reduced by applying a multilevel method instead of a single-level method. An analysis on this point was carried out in [24] for a uniformly bounded diffusion coefficient and bounded parameter space Γ. The results given in [24, p. 8] will be cited below. They are modified for a discretization and for a Galerkin error linear in the mesh width h_k being the relevant cases here. For the more general result and the proof the interested reader is referred to [24, pp. 8–10].

Let γ and C_c be positive constants independent of the mesh width h_k such that the cost C_k to compute $u_{h_k} - u_{h_{k-1}}$ at a sample point is bounded by $C_c h_k^{-\gamma}$ with $k \in \mathbb{N}$, i.e., $C_k \leq C_c h_k^{-\gamma}$.
The cost of the multilevel method is given by

$$\sum_{k=0}^{K} \eta_{K-k} C_k,$$

where the cost in order to compute $u_{h_k} - u_{h_{k-1}}$ is multiplied by the number of collocation points according to this level k, and all levels are summed up. Further, let for the spatial discretization error

$$\|u - u_{h_k}\|_{L^2_\rho(\Gamma;H^1_0(D))} \leq C_1 h_k \tag{5.7}$$

hold, and let ζ be a function of the space $C^0_\sigma(\Gamma; W_{h_k}(D)) \subset L^2_\rho(\Gamma; H^1_0(D))$ to $\mathbb{R}$ with

$$\begin{aligned} \|u_{h_k} - \mathcal{A}_{\eta_k} u_{h_{k-1}}\|_{L^2_\rho(\Gamma;H^1_0(D))} &\leq C_2 \eta_k^{-\mu} \zeta(u_{h_k}) \\ \zeta(u_{h_k}) &\leq C_3 h_0 \\ \zeta(u_{h_k} - u_{h_{k-1}}) &\leq C_3 h_k, \end{aligned} \tag{5.8}$$

where $\mu > 0$. Note that in the previous example, Example 5.12, it was $\zeta(u_{h_k}) = \max_{\mathbf{z}\in\Sigma(\Gamma;\tau)} \sigma(\operatorname{Re}\mathbf{z}) \|u_{h_k}(\mathbf{z})\|_{H^1_0(D)}$. With (5.8), for any $\epsilon < e^{-1}$ there exists an integer K such that the cost to achieve $\|u - u_K^{ML}\|_{L^2_\rho(\Gamma;H^1_0(D))} \leq \epsilon$ is bounded by

$$C_\epsilon^{ML} \leq C \begin{cases} \epsilon^{-\frac{1}{\mu}} & \text{if } 1 > \mu\gamma \\ \epsilon^{-\frac{1}{\mu}} |\log \epsilon|^{1+\frac{1}{\mu}} & \text{if } 1 = \mu\gamma \\ \epsilon^{-\frac{1}{\mu} - \frac{\gamma\mu - 1}{\mu}} & \text{if } 1 < \mu\gamma. \end{cases}$$

For the single-level method, in order to derive the same error estimate, i.e., $\|u - u_{h,\eta}\|_{L^2_\rho(\Gamma;H^1_0(D))} \leq \epsilon$, where the spatial discretization and the stochastic error are of the same size, the computational cost is given by

$$C_\epsilon \approx \epsilon^{-\frac{1}{\mu}-\gamma}.$$

The savings when applying the multilevel method in comparison to the single-level method can be obtained by dividing the multilevel costs by the single level cost yielding

$$\frac{C^{ML}_\epsilon}{C^{SL}_\epsilon} \approx C \begin{cases} \epsilon^\gamma & \text{if } 1 > \mu\gamma \\ \epsilon^\gamma |\log \epsilon|^{1+\frac{1}{\mu}} & \text{if } 1 = \mu\gamma \\ \epsilon^{\frac{1}{\mu}} & \text{if } 1 < \mu\gamma. \end{cases}$$

Since it is $\epsilon < e^{-1} < 1$, it becomes clear that the multilevel method (as being a kind of sparse grid method) saves costs compared to the single-level method. Therefore, it is preferable to a single-level method if the solution, of course, fulfils the necessary assumptions to derive the error estimates.

It shall be noted that the same analysis applies to the mixed form of the diffusion problem with the assumptions on the solution and spaces adjusted to this case: Instead of $H^1_0(D)$ the space $L^2(D) \times H(\mathrm{div}; D)$, and instead of the solution u of the diffusion problem the solution $(p, \mathbf{u})$ of its mixed form have to be considered in the estimates.

Remark 5.14 – Computational costs by quadrature:
Computational costs by the multilevel quadrature might be derived in a similar manner by changing the assumptions on the functions in the estimates of the stochastic approximation. If the estimates for the approximation are replaced by

$$\begin{aligned} \|\mathbb{E}[u] - \mathbb{E}[u_{h_k}]\|_{H^1_0(D)} &\leq C_1 h_k \\ \|\mathbb{E}[u_{h_k}] - Q_{\eta_k}[u_{h_{k-1}}]\|_{H^1_0(D)} &\leq C_2 \eta_k^{-\mu} \zeta(u_{h_k}) \\ \zeta(u_{h_k}) &\leq C_3 h_0 \\ \zeta(u_{h_k} - u_{h_{k-1}}) &\leq C_3 h_k, \end{aligned}$$

where $\mu > 0$, (cf. equations (5.7) and (5.8)) the same costs for the multilevel quadrature can be derived as given in the previous remark for the multilevel approximation because the upper bounds have the same form as for the approximations treated in the previous remark. $\square$

6 Outlook and Conclusions

Some questions arise from this thesis which are worth noting.

Norms for D-bounded diffusion coefficient It has already been mentioned in Remark 5.13 that by a different definition of the norm the desired error estimate of the multilevel method could be derived. This norm has to be some energy norm incorporating the dependence of the constant on the stochastic domain. As mentioned in Remark 3.22, results with some energy norm have been derived in [1] on existence and uniqueness. It might be interesting in the case of a D-bounded diffusion coefficient to carry out the analysis given in this theses to derive results with constants independent of the random source.

Order of convergence The results in this thesis on the multilevel method only consider a linear convergence order. For higher convergence orders h_k^α with $\alpha > 1$ further conditions on the solution and hence on the data are required.

Choice of the polynomial basis In the collocation method only a global polynomial basis was considered. A local basis instead of a global one has the advantage that a disturbance in the data at some point has only local impact on the solution and hence the error. Lagrange polynomials become unstable for a high number of interpolation points. This can be overcome by using a local basis. Research in this direction has already been undertaken (see [14, pp. 589 – 621] for an overview).

Conclusions The analysis and computation of stochastic PDEs is an active field of research. In the literature, in general, the diffusion coefficient is either assumed to be uniformly bounded or bounded by random variables. An overview of analytical results for both cases is given in this thesis which should facilitate the comparison of their assumptions on the data. As a possibility to approximate the solution of the diffusion problem and its mixed form, the collocation method – as one popular method for the numerical computation of stochastic PDEs – was introduced. If the main interest consists in computing the expected value of the solution, the connection to quadrature indicates that it might be worth only applying a “simple” quadrature rule. The use of sparse grids – for single-level methods as well as multilevel methods – reduces the computational costs significantly. Research in this direction will certainly be continued.

A Derivation of the estimate in Lemma 4.8

The aim is to show that the estimate given in [22, p. 2331] is still valid. The authors of [22] do not perform the proof of the estimate given in Lemma 4.8, but they mention on page 2330 that the proof with Gaussian abscissae is almost the same as with Clenshaw-Curtis abscissae. The difference of the estimates with Clenshaw-Curtis and Gaussian abscissae are due to the one-dimensional estimates which have to be applied in the proof, and consists in the constants and a factor i; compare the estimates given on pages 2323 and 2330. The crucial step, where the analyticity of the solution is important, is for the Clenshaw-Curtis abscissae given in [22, p. 2323 – 2325, Lemma 3.4]. As this step is except for the estimates of the one-dimensional results for the Gaussian abscissae identical, the reader should refer to [22, p. 2323 – 2325, Lemma 3.4] to compare the steps.

Note that here, the focus is on showing that the estimate given in Lemma 3.4 in [22] are valid (except for the constant) when assuming the analyticity in all directions simultaneously, while the rest of the proof for deriving the final estimate stays the same and is therefore not given here. In [22] a final estimate of the form

$$\left\|(I - \mathcal{A}^{X(\ell,N)})[\mathbf{w}_h]\right\|_{L^2_\rho(\Gamma;W(D))} \leq \sum_{k=1}^{N} R(\ell,k) \tag{A.1}$$

is sought, where $R(\ell,k) = \frac{1}{2}\sum_{\mathbf{i}\in\tilde{X}(\ell,k)} (2C)^k e^{-\sigma\sum_{n=1}^{k} 2^{i_n-1}}$. This estimate will be shown to hold with the revised assumptions on the solution. According to page 2324 in [22] it holds

$$\begin{aligned} &\left\|(I - \mathcal{A}^{X(\ell,N)})[\mathbf{w}_h]\right\|_{L^2_\rho(\Gamma;W(D))} \\ &= \left\|\sum_{k=2}^{N}\left(\tilde{R}(\ell,k)\bigotimes_{n=k+1}^{N} I_1^{(n)}\right)[\mathbf{w}_h] + \left(I_1^{(1)} - \mathcal{A}^{X(\ell,1)}\right)\bigotimes_{n=2}^{N} I_1^{(n)}[\mathbf{w}_h]\right\|_{L^2_\rho(\Gamma;W(D))}, \end{aligned} \tag{A.2}$$

where $I_1^{(n)} : \Gamma_n \to \Gamma_n$ is the one-dimensional identity operator. The term $\tilde{R}(\ell,k)$ is defined as

$$\tilde{R}(\ell,k) = \sum_{\mathbf{i}\in X(\ell,k-1)} \bigotimes_{n=1}^{k-1} \Delta^{i_n} \otimes (I_1^{(k)} - \mathcal{U}^{\hat{i}_k}), \tag{A.3}$$

with the index $\hat{i}_k = 1 + \ell - \sum_{n=1}^{k-1}(i_k - 1)$ such that $(i_1, \ldots, i_{k-1}, \hat{i}_k)$ is in the set $\tilde{X}(\ell,k)$, for $k \in \{2, \ldots, N\}$. For more details see [22, p. 2324].

The term (A.2) will be estimated using the analyticity of the function $\mathbf{w}_h$ in all N directions. Recall that the solution $\mathbf{w}_h$ is assumed to factorize, i.e., $\mathbf{w}_h = \mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,N}$. In a slightly different context in [3, p. 277] results have been deduced assuming a factorizing function like this.

Step 1: Triangle inequality on (A.2)

$$\begin{aligned}
&\left\|\left(I - \mathcal{A}^{X(\ell,N)}\right)[\mathbf{w}_h]\right\|_{L^2_\rho(\Gamma;W(D))} \\
&\leq \sum_{k=2}^{N} \left\| \left(\tilde{R}(\ell,k) \bigotimes_{n=k+1}^{N} I_1^{(n)} \right) [\mathbf{w}_h] \right\|_{L^2_\rho(\Gamma;W(D))} \\
&\quad + \left\| \left(\left(I_1^{(1)} - \mathcal{A}^{X(\ell,1)}\right) \bigotimes_{n=2}^{N} I_1^{(n)} \right) [\mathbf{w}_h] \right\|_{L^2_\rho(\Gamma;W(D))}
\end{aligned} \tag{A.4}$$

Step 2: Cross norm property

In [3, p. 278] the cross norm property for functions was applied. Here the one for continuous linear operators will be used. Since the operators $\tilde{R}(\ell,k)\bigotimes_{n=k+1}^{N} I_1^{(n)}$ and $(I_1^{(1)} - \mathcal{A}^{X(\ell,1)})\bigotimes_{n=2}^{N} I_1^{(n)}$ are by definition of their components continuous linear maps in $(\Gamma \to W(D))$, the property

$$\begin{aligned}
&\left(\tilde{R}(\ell,k) \bigotimes_{n=k+1}^{N} I_1^{(n)} \right) [\mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,N}] \\
&= \tilde{R}(\ell,k)[\mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,k}] \bigotimes_{n=k+1}^{N} I_1^{(n)}[\mathbf{w}_{h,n}]
\end{aligned} \tag{A.5}$$

applies. In the same way it follows

$$\begin{aligned}
&\left((I_1^{(1)} - \mathcal{A}^{X(\ell,1)}) \bigotimes_{n=2}^{N} I_1^{(n)} \right) [\mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,N}] \\
&= (I_1^{(1)} - \mathcal{A}^{X(\ell,1)})[\mathbf{w}_{h,1}] \bigotimes_{n=2}^{N} I_1^{(n)}[\mathbf{w}_{h,n}].
\end{aligned} \tag{A.6}$$

Step 3a: Estimation of (A.5) **in its corresponding norm**

Since $L^2_\rho(\Gamma; W(D))$ is a Hilbert space if $W(D)$ is a Hilbert space (the scalar product inducing the norm is given by $\langle f,g\rangle = \int_\Gamma \langle f,g\rangle_{W(D)}\mathrm{d}\mathbf{y}$), it holds

$$
\begin{aligned}
&\Big\|\tilde{R}(\ell,k)[\mathbf{w}_{h,1}\otimes\cdots\otimes\mathbf{w}_{h,k}]\bigotimes_{n=k+1}^{N} I_1^{(n)}[\mathbf{w}_{h,n}]\Big\|_{L^2_\rho(\Gamma;W(D))}\\
=&\big\|\tilde{R}(\ell,k)[\mathbf{w}_{h,1}\otimes\cdots\otimes\mathbf{w}_{h,k}]\big\|_{L^2_{\rho_1\cdots\rho_k}(\Gamma_1\times\cdots\times\Gamma_k;W(D))}\\
&\cdot\prod_{n=k+1}^{N}\big\|I_1^{(n)}[\mathbf{w}_{h,n}]\big\|_{L^2_{\rho_n}(\Gamma_n;W(D))}.
\end{aligned}
\tag{A.7}
$$

Inserting (A.3) in the first term of (A.7) it is obtained by again applying the cross norm property

$$
\begin{aligned}
&\big\|\tilde{R}(\ell,k)[\mathbf{w}_{h,1}\otimes\cdots\otimes\mathbf{w}_{h,k}]\big\|_{L^2_{\rho_1\cdots\rho_k}(\Gamma_1\times\cdots\times\Gamma_k;W(D))}\\
\overset{(A.3)}{\leq}&\sum_{\mathbf{i}\in X(\ell,k-1)}\Big\|\bigotimes_{n=1}^{k-1}\Delta^{i_n}\otimes(I_1^{(k)}-\mathcal{U}^{\hat{i}_k})[\mathbf{w}_{h,1}\otimes\cdots\otimes\mathbf{w}_{h,k}]\Big\|_{L^2_{\rho_1\cdots\rho_k}(\Gamma_1\times\cdots\times\Gamma_k;W(D))}\\
=&\sum_{\mathbf{i}\in X(\ell,k-1)}\Big\|\bigotimes_{n=1}^{k-1}\Delta^{i_n}[\mathbf{w}_{h,1}\otimes\cdots\otimes\mathbf{w}_{h,k-1}]\Big\|_{L^2_{\rho_1\cdots\rho_{k-1}}(\Gamma_1\times\cdots\times\Gamma_{k-1};W(D))}\\
&\cdot\big\|(I_1^{(k)}-\mathcal{U}^{\hat{i}_k})[\mathbf{w}_{h,k}]\big\|_{L^2_{\rho_k}(\Gamma_k;W(D))}\\
\leq&\sum_{\mathbf{i}\in X(\ell,k-1)}\prod_{n=1}^{k-1}\big\|\Delta^{i_n}[\mathbf{w}_{h,n}]\big\|_{L^2_{\rho_n}(\Gamma_n;W(D))}\cdot\big\|(I_1^{(k)}-\mathcal{U}^{\hat{i}_k})[\mathbf{w}_{h,k}]\big\|_{L^2_{\rho_k}(\Gamma_k;W(D))}\\
\leq&\ R(\ell,k)\max_{z_1\in\Sigma(\Gamma_1;\tau_1)}\|\mathbf{w}_{h,1}(z_1)\|_{W(D)}\cdots\max_{z_k\in\Sigma(\Gamma_k;\tau_k)}\|\mathbf{w}_{h,k}(z_k)\|_{W(D)}.
\end{aligned}
\tag{A.8}
$$

The last step follows using the one-dimensional results given in [22, p. 2330] in analogue to the proof in [22]. For the second term of (A.7) it holds

$$
\begin{aligned}
\big\|I_1^{(n)}[\mathbf{w}_{h,n}]\big\|_{L^2_{\rho_n}(\Gamma_n;W(D))}&=\left(\int_{\Gamma_n}\|\mathbf{w}_{h,n}(y_n)\|^2_{W(D)}\rho_n(y_n)\mathrm{d}y_n\right)^{1/2}\\
&\leq\max_{z_n\in\Sigma(\Gamma_n;\tau_n)}\|\mathbf{w}_{h,n}(z_n)\|_{W(D)}\cdot\sqrt{\int_{\Gamma_n}\rho_n(y_n)\mathrm{d}y_n}.
\end{aligned}
\tag{A.9}
$$

Inserting (A.8) and (A.9) in (A.7) it follows with

$$
\max_{z_1\in\Sigma(\Gamma_1;\tau_1)}\|\mathbf{w}_{h,1}(z_1)\|_{W(D)}\cdots\max_{z_N\in\Sigma(\Gamma_N;\tau_N)}\|\mathbf{w}_{h,N}(z_N)\|_{W(D)}=\max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})}\|\mathbf{w}_h(\mathbf{z})\|_{W(D)}
$$

$$\left\| \tilde{R}(\ell,k)[\mathbf{w}_{h,1} \otimes \cdots \otimes \mathbf{w}_{h,k}] \bigotimes_{n=k+1}^{N} I_1^{(n)}[\mathbf{w}_{h,n}] \right\|_{L^2_\rho(\Gamma;W(D))}$$
$$\leq R(\ell,k) \max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)} \prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y_n)\mathrm{d}y_n}. \tag{A.10}$$

Step 3b: Estimation of (A.6) **in its corresponding norm**
By the same proceeding the following estimate is valid:

$$\left\| (I_1^{(1)} - \mathcal{A}^{X(\ell,1)})[\mathbf{w}_{h,1}] \bigotimes_{n=2}^{N} I_1^{(n)}[\mathbf{w}_{h,n}] \right\|_{L^2_\rho(\Gamma;W(D))}$$
$$\leq R(\ell,1) \max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)} \prod_{n=2}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y_n)\mathrm{d}y_n}. \tag{A.11}$$

Step 4: Final estimate
With the last two estimates (A.10) and (A.11) it follows for (A.2)

$$\left\| (I - \mathcal{A}^{X(\ell,N)})[\mathbf{w}_h] \right\|_{L^2_\rho(\Gamma;W(D))}$$
$$\leq \sum_{k=1}^{N} R(\ell,k) \max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)} \prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y_n)\mathrm{d}y_n}.$$

Note that this is except for the factor $\prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y_n)\mathrm{d}y_n}$, which can be incorporated in the constant C of $R(\ell,k)$, see below, and the factor $\max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)}$ the estimate given in the lemma in [22], see (A.1). There, the maximal value (but in one direction) was assumed to equal the constant 1. Therefore, it does not appear in the estimate given there. Notice as well that the analyticity is required in all directions simultaneously as the term $\max_{\mathbf{z}\in\Sigma(\Gamma;\boldsymbol{\tau})} \|\mathbf{w}_h(\mathbf{z})\|_{W(D)}$ indicates.
The constant C is given on p. 2330 in [22] by $C = \frac{8}{e^{2\sigma}-1}\sqrt{\int_{\Gamma_k} \rho_k(y)\mathrm{d}y}$. Multiplying $R(\ell,k)$ with $\prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y_n)\mathrm{d}y_n}$ gives another constant, C_1, with $C_1 = \frac{8}{e^{2\sigma}-1}\sqrt{\int_{\Gamma_k} \rho_k(y)\mathrm{d}y}\left(\prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y)\mathrm{d}y}\right)^{1/k}$. This constant is dependent on k. It gets independent of k by taking the maximal value of $k \in \{1,\ldots,N\}$:

$$C_1 \leq \frac{8}{e^{2\sigma}-1} \max_{k=1,\ldots,N} \sqrt{\int_{\Gamma_k} \rho_k(y)\mathrm{d}y}\left(\prod_{n=k+1}^{N} \sqrt{\int_{\Gamma_n} \rho_n(y)\mathrm{d}y}\right)^{1/k}.$$

Having obtained this estimate, the rest of the proof follows analogously to [22]. Hence, the estimate given in Theorem 4.8 with the modified assumptions holds.

B Functional Analysis

Definition B.1 – Lebesgue spaces:
Let $p \in [1, \infty)$. The Lebesgue space

$$L^p(D) = \left\{ v : D \to \mathbb{R} \text{ measurable} : \int_D |v(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} < \infty \right\}$$

consists of Lebesgue integrable functions to the power p with finite norm

$$\|v\|_{L^p(D)} = \left(\int_D |v(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} < \infty \right)^{1/p}.$$

For $p = \infty$, the Lebesgue space is defined as

$$L^\infty(D) = \left\{ v : D \to \mathbb{R} \text{ measurable} : \operatorname*{ess\,sup}_{\mathbf{x} \in D} |v(\mathbf{x})| < \infty \right\}$$

equipped with norm

$$\|v\|_{L^\infty(D)} = \operatorname*{ess\,sup}_{\mathbf{x} \in D} |v(\mathbf{x})|.$$

□

Lemma B.2 – Hölder and Cauchy-Schwarz inequality:
Let $p,\ q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p_\rho(D)$ and $g \in L^q_\rho(D)$, where ρ is a density on D, then $fg \in L^1_\rho(D)$ and it holds

$$\|fg\|_{L^1_\rho(D)} \leq \|f\|_{L^p_\rho(D)} \|g\|_{L^q_\rho(D)}.$$

For $p = q = 2$ the so called Cauchy-Schwarz inquality is given as a particular case by

$$\|fg\|_{L^1_\rho(D)} \leq \|f\|_{L^2_\rho(D)} \|g\|_{L^2_\rho(D)}.$$

□

Definition B.3 – Sobolev-spaces:
Let $p \in [1, \infty]$, $k \in \mathbb{N}$. The Sobolev space $W^{k,p}(D)$ is the space

$$W^{k,p}(D) = \{ v \in L^p(D) : D^{\boldsymbol{\alpha}} v \in L^p(D) \text{ with } |\boldsymbol{\alpha}| \leq k \},$$

of functions in $L^p(D)$ whose weak derivatives $D^{\boldsymbol{\alpha}}$ of order $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}| \leq k$ are elements of $L^p(D)$.

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d), \quad d = \dim D$$

is a multi-index with

$$|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_d.$$

The Sobolev space $W^{k,p}(D)$ *is equipped with the norm*

$$\|v\|_{W^{k,p}} = \sum_{|\boldsymbol{\alpha}|\leq k} \|D^{\boldsymbol{\alpha}} v\|_{L^p(D)}.$$

□

Remark B.4:
A special case is the Sobolev space for $p = 2$ denoted by $H^k(D) = W^{k,2}(D)$. This space is a Hilbert space with inner product

$$(u, v)_{H^k(D)} = \sum_{|\boldsymbol{\alpha}|\leq k} \int_D D^{\boldsymbol{\alpha}} u(\mathbf{x}) D^{\boldsymbol{\alpha}} v(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

and norm

$$\|v\|_{H^k(D)} = (u, v)^{1/2}_{H^k(D)}.$$

Further, the space $H_0^1(D)$ is of particular interest here. It is defined by

$$\begin{aligned} H_0^1(D) &= \{v \in H^1(D) : v|_{\partial(D)} = 0\} \\ &= \{v \in L^2(D) : \nabla v \in L^2(D), v|_{\partial(D)} = 0\}. \end{aligned}$$

The following norm (which is equivalent to the "standard" norm on $H_0^1(D)$) will be considered throughout the thesis

$$\|v\|_{H_0^1(D)} = \|\nabla v\|_{L^2(D)}.$$

□

The following definition is necessary to define Sobolev spaces with negative exponents.

Definition B.5 – The space $\boldsymbol{C_0^\infty(D)}$:
The space

$$C_0^\infty(D) = \{v : v \in C^\infty,\ \mathrm{supp}(v) \subset D\}$$

consists of infinitely often differentiable functions whose support in D *is compact.* □

Definition B.6 – Sobolev spaces with negative exponents:
Let $p \in (1, \infty)$, q such that $p^{-1} + q^{-1} = 1$ and $k \in \mathbb{N}$. The space

$$W^{-k,q}(D) = \{\phi \in (C_0^\infty(D))^* : \|\phi\|_{W^{-k,q}(D)}\} < \infty,$$

where the norm is defined as

$$\|\phi\|_{W^{-k,q}(D)} = \sup_{v \in C_0^\infty(D)} \frac{\langle \phi, v\rangle}{\|v\|_{W^{k,p}(D)}},$$

is a Sobolev space with negative exponent $-k$. □

Remark B.7:
It is $W^{-k,q}(D) = (W^{k,p}(D))^*$, i.e., $W^{-k,q}(D)$ can be identified with the dual space of $W^{k,p}(D)$. □

Lemma B.8 – Lemma of Lax-Milgram:
Let V be a Hilbert space and $B : V \times V \to \mathbb{R}$ be a bounded and coercive bilinear form, i.e., it holds

$$|B(u, v)| \leq M\|u\|_V \|v\|_V \ \forall u, v \in V,\ M > 0$$

and

$$B(u, u) \geq m\|u\|_V^2 \ \forall u \in V,\ m > 0,$$

where the constants m, M are independent of u and v.
Then, for each bounded functional $f \in V^*$, where V^* is the dual space of V, there exists a unique $u \in V$ such that $B(u, v) = f(v)\ \forall v \in V$. □

Lemma B.9 – Poincaré inequality:
Let $D \subset \mathbb{R}^d$ be a bounded domain with Lipschitz-boundary ∂D. Then it holds

$$\|v\|_{L^2(D)} \leq C_P \|\nabla v\|_{L^2(D)} \quad \forall v \in H_0^1(D)$$

with a constant C_P only depending on the diameter of the domain D. □

Theorem B.10 – Cf. Theorem 1.1, p. 42 in [6] and p. 2059 in [5]:

Let $A(\cdot,\cdot)$ be a continuous linear form on $V \times V$ and $B(\cdot,\cdot)$ be a continuous linear form on $V \times Q$, where V and Q are Hilbert spaces. Assume that there exists $k_0 > 0$ such that the inf-sup-condition

$$\sup_{\mathbf{v}\in V} \frac{B(v,q)}{\|\mathbf{v}\|_V} \geq k_0 \|q\|_Q \quad \forall q \in Q$$

holds. Further, let

$$A(\mathbf{v},\mathbf{v}) \geq \alpha_0 \|\mathbf{v}\|_V^2 \quad \forall \mathbf{v} \in V^0 = \{\mathbf{v} \in V : B(\mathbf{v},q) = 0 \;\; \forall q \in Q\}.$$

Then, the problem

$$\begin{aligned} A(\mathbf{u},\mathbf{v}) + B(\mathbf{v},p) &= l(\mathbf{v}) \quad \forall \mathbf{v} \in V \\ B(\mathbf{u},q) &= h(q) \quad \forall q \in Q, \end{aligned}$$

where $l : V \to \mathbb{R}$ and $h : Q \to \mathbb{R}$ are bounded linear functionals, admits a unique solution $(p,\mathbf{u}) \in Q \times V$. With $l(\mathbf{v}) = \int_{\partial D} g(\mathbf{x})\mathbf{v}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})\,\mathrm{d}\mathbf{x}$ and $h(q) = \int_D f(\mathbf{x})q(\mathbf{x})\,\mathrm{d}\mathbf{x}$ the following bounds hold

$$\begin{aligned} \|p\|_Q &\leq \left(\frac{\|A\|}{\alpha_0} + 1\right)\frac{1}{k_0}\|g\|_{V^*} + \frac{\|A\|}{k_0^2}\left(\frac{\|A\|}{\alpha_0} + 1\right)\|f\|_{Q^*} \\ \|\mathbf{u}\|_V &\leq \frac{1}{\alpha_0}\|g\|_{V^*} + \left(\frac{\|A\|}{\alpha_0} + 1\right)\frac{1}{k_0}\|f\|_{Q^*}. \end{aligned}$$

□

C Analytic Functions

Some definitions and results on analytic functions are given.

Definition C.1 – Complex differentiability in several variables:
Let $W \subset \mathbb{C}^n$ be an open domain, $w \in W$. A function $h : W \to \mathbb{C}$ is called complex differentiable in the point w if there exists a map $\Delta : W \to \mathbb{C}^n$ such that

1. *Δ is continuous in w and*
2. *$h(z) = h(w) + \Delta(z) \cdot (z - w)$ for $z \in W$.*

For $\Delta = (\Delta_1, \ldots, \Delta_n)$ it is

$$h(z_1, \ldots, z_n) = h(w_1, \ldots, w_n) + \sum_{\nu=1}^{n} \Delta_\nu(z_1, \ldots, z_n) \cdot (z_\nu - w_\nu)$$

If h is complex differentiable, $\Delta(w)$ is called derivative of h in w, and the numbers $\Delta_\nu(w) = e_\nu \cdot \Delta(w) = \partial_\nu h(w) = \frac{\partial h}{\partial z_\nu}(w)$ are the uniquely defined partial derivatives of h in w. □

Theorem C.2 – Theorem of Osgood:
Let $W \in \mathbb{C}^n$ be open. Then, for $h : W \to \mathbb{C}$ the following statements are equivalent

1. *h is holomorphic*
2. *h is complex differentiable*
3. *h is weakly holomorphic.*

(Therefore, often the notion of holomorphic/analytic and complex differentiable are used synonymously.) □

Remark C.3:
A theorem with fewer requirements (i.e., h is partial differentiable only and not weakly holomorphic) follows by the Theorem of Hartogs. □

Theorem C.4 – Theorem of Hartogs:
Let $W \subset \mathbb{C}^n$ be open, $h : W \to \mathbb{C}$. If $z \mapsto h(z_1, \ldots, z_{j-1}, z, z_{j+1}, \ldots, z_N)$ is an analytic map in z for each fixed set of $(z_1, \ldots, z_{j-1}, z, z_{j+1}, \ldots, z_N)$, where $z \in \{z \in \mathbb{C} : (z_1, \ldots, z_{j-1}, z, z_{j+1}, \ldots, z_N) \in D\}$, then the function h is continuous on D. □

D Notations

Notation	Description	Page
a	diffusion coefficient	3
$\mathcal{A}$	approximation operator	52
$\mathcal{A}_{\eta_k}$	approximation operator in dependence of the number η_k of collocation points	76
$\mathcal{A}_N^{\mathbf{i}}$	tensor product approximation operator	54
$\mathcal{A}^{X(\ell,N)}$	Smolyak approximation operator on the set $X(\ell, N)$	56
$C^0_\sigma(\Gamma; W(D))$	space of continuous functions from Γ to $W(D)$ with norm weighted by σ	17
d	$\dim D$	3
D	spatial domain, $\mathbf{x} \in D$, with boundary ∂D	3
f	force term of the diffusion problem & its mixed form	4
$\mathbf{f}$	force term of the general elliptic problem	3
Γ	parameter space, $\mathbf{y} \in \Gamma$ of dimension N	8
Γ_k	parameter space in direction k only	8
$H(\mathrm{div}; D)$	space of $L^2(D)$ functions $\mathbf{w}$ with $\nabla \cdot \mathbf{w} \in L^2(D)$	24
$L^2_\rho(\Gamma; W(D))$	L^2-space of functions from Γ to $W(D)$ weighted by ρ	10
$(p, \mathbf{u})$	solution of the diffusion problem's mixed form	4
Q	quadrature formula	66
Q_{η_k}	quadrature formula in dependence of the number η_k of quadrature points	76
$Q_N^{\mathbf{i}}$	tensor product of N quadrature formulae	60
$Q^{X(\ell,N)}$	Smolyak quadrature formula on the set $X(\ell, N)$	61
ρ	density function from Γ to $\mathbb{R}_+$	8
σ	weight function from Γ to $\mathbb{R}_+$	16
$\Sigma(\Gamma; \boldsymbol{\tau})$	region of the complex plane $\mathbb{C}^N$ to which the solution has an analytic continuation	71
Σ_D	subset of $\mathbb{C}^N$, defined by the diffusion coefficient which is bounded in D	28
Σ_U	subset of $\mathbb{C}^N$, defined by the uniformly bounded diffusion coefficient	19
u	solution of the diffusion problem	4
$\mathbf{w}$	solution of the general elliptic problem	3
$\mathbf{w}_h$	discretization of $\mathbf{w}$ in spatial variable	51
$\mathbf{w}_{h,\eta}$	discretization of $\mathbf{w}$ in spatial and stochastic variable	52
$\mathcal{W}$	$H^1_0 \cap H^2$	42
$W(D)$	banach space of functions from D to $\mathbb{R}^n$	4
Ω	probability space, $\omega \in \Omega$	3

References

[1] Ivo Babuška, Fabio Nobile and Raúl Tempone (2007): *A stochastic collocation method for elliptic partial differential equations with random input data.* SIAM Journal on Numerical Analysis, 45(3), pp. 1005-1034.

[2] Andrea Barth, Christoph Schwab and Nathaniel Zollinger (2011): *Multilevel Monte Carlo Finite Element method for elliptic PDEs with stochastic coefficients.* Numerische Mathematik, 119(1), pp.123-161.

[3] Volker Barthelmann, Erich Novak and Klaus Ritter (2000): *High dimensional polynomial interpolation on sparse grids.* Advances in Computational Mathematics, 12, pp. 273-288.

[4] Joakim Beck, Fabio Nobile, Lorenzo Tamellini and Raúl Tempone (2011): *On the optimal approximation of stochastic PDEs by Galerkin and Collocation methods.* MOX-Report No. 23/2011.

[5] Alexei Bespalov, Catherine E. Powell and David Silvester (2012): *A priori analysis of stochastic Galerkin mixed approximations of elliptic PDEs with random data.* Siam Jounal on Numerical Analysis, 50(4), pp. 2039- 2063.

[6] Franco Brezzi and Michel Fortin (1991): *Mixed and hybrid finite element methods.* New York: Springer-Verlag.

[7] Albert Cohen, Ronald DeVore and Christoph Schwab (2011): *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs.* Analysis and Applications, 9(1), pp. 11– 47.

[8] Ronald A. DeVore, Hoang A. Tran, Clayton G. Webster and Guannan Zhang (2014): *Analysis of quasi-optimal polynomial approximations for parametrized PDEs with deterministic and stochastic coefficients.* ORNL/TM-2014/468.

[9] Jürgen Elstrodt (2011): *Maß- und Integrationstheorie. 7., korrigierte und aktualisierte Auflage.* Berlin, Heidelberg: Springer-Verlag.

[10] Oliver G. Ernst and Björn Sprungk (2012): *Stochastic Collocation for Elliptic PDEs with random data – the lognormal case.* Submitted paper, available on: `http://www.mathe.tu-freiberg.de/~ernst/PubArchive/sga2012.pdf` [July 27,2015].

[11] Oliver G. Ernst and Björn Sprungk (2014): *Stochastic Collocation for Elliptic PDEs with random data – the lognormal case.* Available on: `https://www.tu-chemnitz.de/mathematik/numa/PubArchive/ernstSprungk2014.pdf` [July 27,2015].

[12] Klaus Fritzsche and Hans Grauert (2002): *From holomorphic functions to complex manifolds.* Graduate texts in mathematics, 213. New York: Springer-Verlag.

[13] Thomas Gerstner and Michael Griebel (1998): *Numerical Integration using Sparse Grids.* Numerical Algorithms, 18, pp. 209-232.

[14] Max D. Gunzburger, Clayton G. Webster and Guannan Zhang (2014): *Stochastic finite element methods for partial differential equations with random input data.* Acta Numerica, 23, pp. 521-650.

[15] Helmut Harbrecht, Michael Peters and Markus Siebenmorgen (2013): *Multilevel accelerated quadrature for PDEs with log-normal distributed random coefficient.* Preprint 2013-18, Mathematisches Institut, Universität Basel, Switzerland.

[16] Helmut Harbrecht, Michael Peters and Markus Siebenmorgen (2011): *On multilevel quadrature for elliptic stochastic partial differential equations.* Preprint 2011-01, Mathematisches Institut, Universität Basel, Switzerland.

[17] Volker John (2013): *Numerical Mathematics III – Partial Differential Equations.* Lecture Notes, Berlin.

[18] Volker John (2014): *Numerical methods for incompressible flow problems I.* Lecture Notes, Berlin.

[19] Steven G. Krantz (1982): *Function theory of several complex variables.* John Wiley & Sons, Inc..

[20] Antje Mugler, Hans-Jörg Starkloff (2011): *On Elliptic Partial Differential Equations with Random Coefficients.* DFG-Schwerpunktprogramm 1324: Extraktion quantifizierbarer Information aus komplexen Systemen, Preprint 79.

[21] Fabio Nobile, Lorenzo Tamellini and Raúl Tempone (2014): *Convergence of quasi-optimal sparse grid approximation of Hilbert-valued functions: application to random elliptic PDEs.* Mathcise Technical Report Nr. 12.2014, EPFL, Lausanne, Switzerland.

[22] F. Nobile, R. Tempone and C. G. Webster (2008a): *A sparse grid stochastic collocation method for partial differential equations with random input data.* SIAM Journal on Numerical Analysis, 46(5), pp. 2309-2345.

[23] F. Nobile, R. Tempone and C. G. Webster (2008b): *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data.* SIAM Journal on Numerical Analysis, 46(5), pp. 2411-2442.

[24] A. L. Teckentrup, P. Jantsch, C. G. Webster and M. Gunzburger (2014): *A multilevel stochastic collocation method for partial differential equations with random input data.* Preprint, arXiv:1404.2647 [math.NA].

[25] Grzegorz W. Wasilkowski and Henryk Woźniakowski (1995): *Explicit Cost Bounds of Algorithms for Multivariate Tensor Product Problems.* Journal of Complexity, 11, pp. 1-56.

[26] Dongbin Xiu and Jan S. Hesthaven (2005): *High-order collocation methods for differential equations with random inputs.* Siam Journal on Scientific Computing, 27(3), pp. 1118-1139.

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Masterarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe.
Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.
Diese Arbeit wurde in gleicher oder ähnlicher Form noch bei keiner anderen Universität als Prüfungsleistung eingereicht und ist auch noch nicht veröffentlicht.

Berlin, 25. 10. 2015