

# Sparse Grid Methods for High-dimensional Problems in Uncertainty Quantification

Uta Seidler

Born 24.08.1996 in Haan, Germany

21.04.2021

Master's Thesis Mathematics

Advisor: Prof. Dr. Michael Griebel

Second Advisor: Prof. Dr. Marc Alexander Schweitzer

INSTITUT FÜR NUMERISCHE SIMULATION

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT DER  
RHEINISCHEN FRIEDRICH-WILHELMS-UNIVERSITÄT BONN



# Acknowledgments

First, I would like to thank Prof. Dr. Michael Griebel for introducing me to this very interesting topic and supervising my thesis. His guidance and suggestions have provided me with new ideas and have been helpful throughout the process. I also want to thank him for the opportunity to work at the Institute for Numerical Simulation. In addition, I would like to thank him, as well as Prof. Dr. Marc Alexander Schweitzer, for evaluating my thesis.

Furthermore, I am grateful to my friends and family for proof-reading parts of the thesis and making valuable remarks. In particular, I want to thank my father for improving my English and special thanks go to my sister for offering her advise whenever needed. Finally, I want to express my gratitude for the love, encouragement, and constant support of my family throughout my studies and life in general.

Thank you!



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Problem Description</b>	<b>5</b>
2.1	Finite-dimensional Stochastic . . . . .	5
2.2	Infinite-dimensional Stochastic . . . . .	6
<b>3</b>	<b>Numerical Methods</b>	<b>7</b>
3.1	Quadrature Methods . . . . .	7
3.1.1	Monte Carlo Sampling . . . . .	8
3.1.2	Univariate Quadrature Rules and Product Quadrature . . . . .	10
3.2	Finite Element Method . . . . .	14
3.3	Finite Noise Truncation . . . . .	15
<b>4</b>	<b>Sparse Grids and the Combination Technique</b>	<b>22</b>
4.1	Sparse Grids . . . . .	23
4.1.1	General Concept . . . . .	24
4.1.2	Mixed Regularity . . . . .	27
4.2	Combination Technique . . . . .	29
4.3	Numerical examples . . . . .	31
<b>5</b>	<b>Anisotropic Sparse Grids</b>	<b>37</b>
5.1	Anisotropic Sparse Grid Method for Quadrature and Spatial Discretization . . . . .	38
5.1.1	Numerical Examples . . . . .	42
5.2	Sparse Grid for Truncation and Quadrature . . . . .	44
5.2.1	Monte Carlo . . . . .	49
5.2.2	Product Quadrature Rule . . . . .	52
5.2.3	Numerical Examples . . . . .	61
<b>6</b>	<b>Dimension-adaptive Combination Technique</b>	<b>66</b>
6.1	Finite-dimensional Parameter Space . . . . .	67
6.2	Infinite-dimensional Parameter Space . . . . .	71
6.3	Numerical Examples . . . . .	75
<b>7</b>	<b>Application to Darcy Flow Problem</b>	<b>85</b>
7.1	Mixed Finite Elements . . . . .	87
7.2	Numerical Examples . . . . .	88
<b>8</b>	<b>Summary and Outlook</b>	<b>94</b>



# 1 Introduction

Mathematical models and computational simulations are widely used in diverse applications to describe systems in natural sciences and engineering. Often, however, the input parameters of the models are uncertain. Uncertainty quantification aims at characterizing these uncertainties, includes them in the mathematical description, and tries to quantify how likely outcomes of the model are.

An example for such a mathematical model that describes some real world application is a partial differential equation (PDE) which is given by

$$\mathcal{L}(\omega)(u(\omega)) = f(\omega) \quad \text{in } D \subset \mathbb{R}^n \quad (1)$$

together with suitable boundary conditions. In this, case the operator  $\mathcal{L}$ , the forcing term  $f$ , and the boundary conditions can be uncertain. This randomness of the input might be expressed in a dependence on a set of random variables, or random fields.

Uncertainty quantification deals with problems in many areas, such as physics, chemistry, economics, meteorology. Examples, where the model is of the form (1), are the deformation of solid objects modeled by linear elasticity, where the properties of the material are considered as random variables, or the description of groundwater flow, where the permeability of the rock or sand is represented by a random field.

In general, the uncertainty can stem from different sources. Often a distinction is made between epistemic uncertainty and aleatoric uncertainty, although the classification is not always clear. Epistemic uncertainties arise due to a lack of knowledge, for example missing information of physical properties of a medium, or simplifying model assumptions, whereas aleatoric uncertainties cannot be reduced by additional knowledge. They are inherent to the problem, for example the location and magnitude of earthquakes ([1]).

For any given model, the uncertainty in the input parameters will propagate through the model such that the output will also be random. Hence, the solution  $u$  of (1) is itself a random function and  $u(\omega)$  is just one realization. In many applications, not the random function  $u(\omega)$  is of interest, but rather the value of some function applied to  $u$ . These values are known as quantities of interests. For example, one might be interested in the statistics of  $u$ , meaning the expectation, variance, or higher moments. Other examples include the probability for  $u$  exceeding a certain value or the average of the function in a subdomain of  $D$ . The quantities of interest are often specific to the application (see [1] and [2] for more examples).

Uncertainty quantification is a broad area and includes many aspects. While forward uncertainty quantification analyzes, given the distribution of the randomness in the input parameters, how the uncertainty propagates through the model, inverse uncertainty quantification problems are interested in improving the uncertainty description of the input data based on available measurements of the output of the model. Other aspects include global sensitivity analysis and

determine which random input variables have the largest influence on the output. Further fields in uncertainty quantification are optimization and design of experiments, model validation, model calibration (see [3, 1]). In this thesis, we will focus on the propagation of the uncertain parameters through the model and discuss how to compute the quantities of interest. For this, we assume that the distribution of the input parameters are given.

In practice, the models investigated are complex and computationally challenging to evaluate, in particular when the dimensionality of the problem is large. Hence, there is a typical trade-off between the computational cost and accuracy. In this thesis, we will address how to obtain a result with high accuracy while keeping the cost low when various approximations are applied.

We will focus on the elliptic PDE

$$-\operatorname{div}(a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) = f(\mathbf{x}) \quad \text{in } D$$

as a model problem, where the diffusion coefficient  $a(\mathbf{x}, \omega)$  is subject to uncertainty. Either it depends on a finite set of random variables or it is given by a random field which is in general infinite-dimensional. This equation is often studied, for example, in groundwater flow problems where the permeability of the medium varies randomly from one point in the spatial domain  $D$  to another and the diffusion coefficient is therefore described by a random field with a given covariance function. As quantities of interest, we will compute the expectation  $\mathbb{E}[u(\mathbf{x})]$  and second moment  $\mathbb{E}[u^2(\mathbf{x})]$ .

In order to compute these quantities of interest, several numerical approximations need to be applied. First, in the case where the diffusion coefficient is given by a random field, the random field has to be replaced by a finite-dimensional approximation. A widely used approach is the Karhunen-Loeve expansion (see, e.g., [4]). Given the covariance function and mean of a random field on a compact domain with bounded variance and continuous covariance, the random field is expanded into a series of random variables which are uncorrelated and sometimes independent. The coefficients of the expansion exhibit a certain decay. Thus, the influence of the random variables in the model decreases. A finite-dimensional approximation to the random field is obtained by the truncation of the series. It is then convenient to parametrize the randomness of the input by replacing the random variables by parameters  $\mathbf{y}$  which are defined on the image of the random variables. Depending on the decay rate of the coefficients which is related to the smoothness of the covariance function, this can easily lead to a high-dimensional parametric problem.

Second, the evaluation of  $\mathbb{E}[u(\mathbf{x})]$  and  $\mathbb{E}[u^2(\mathbf{x})]$  requires solving a high-dimensional Bochner integration problem. We will focus on collocation and sampling methods. A classical approach is the Monte Carlo method which is based on random sampling and averaging of the outcomes. It has a major advantage in being independent of the stochastic dimension but converges only with a rate of  $1/2$ . If the function under consideration is to some degree differentiable, other quadrature methods can provide better convergence results. In this case, the quadrature

points are deterministic and one-dimensional quadrature methods can be combined to build a product quadrature.

In either case, the quadrature method requires the repeated evaluation of the integrand, which means that the PDE must be solved for many different values of the parameter  $\mathbf{y}$ . For that, a third approximation needs to be considered. For each point  $\mathbf{y}$ , the deterministic PDE is solved by a discretization in the spatial variable, for example, by applying a finite element method.

In order to have a good approximation to the quantities of interests, the truncation level, spatial and stochastic discretizations have to be chosen such that the errors are small. This leads to a large computational effort. A small error requires a large truncation level, leading to a high-dimensional integration problem which needs to be approximated with many quadrature points. Moreover, for each quadrature point the PDE must be solved with a high accuracy. This exponential growth of computational complexity with dimension, due to combining different methods, is known as the curse of dimensionality.

Throughout this thesis, we consider the sparse grid method in order to break the curse of dimensionality. This method balances the different numerical approximations by combining high with low discretization levels of each numerical method. In particular, we concentrate on the representation of the sparse grid in form of the combination technique. Provided the solution possesses some kind of mixed regularity, it allows reducing the computational cost, while almost preserving the accuracy of a single-level combination of the numerical methods.

In general, two different approaches can be pursued. On the one hand, a sparse grid is constructed based on a-priori theoretical estimates that hold for a certain class of problems. For example, the sparse grid is constructed such that convergence rates can be assured for problems where the solution  $u$  satisfies mixed regularity conditions in terms of  $\mathbf{y}$  and  $\mathbf{x}$ . The a-priori approach relies on knowledge about properties of the solution. On the other hand, an a-posteriori approach can be applied, where the sparse grid structure is discovered in the course of the algorithm and is based on suitable indicators. While no theoretical estimates are required, the sparse grid obtained is specific to the problem investigated.

In the first part of this thesis, we will take the a-priori perspective and discuss how to apply a sparse grid regarding different numerical approximations in the process of computing the quantities of interest. To begin with, we use a sparse grid within a product quadrature to reduce the computational complexity of the high-dimensional integration. In addition, we include the finite element approximation while keeping the truncation level fixed and consider a sparse grid between the spatial and stochastic discretization. We adapt the work of [5] and [6] to construct a sparse grid such that the finite element and quadrature methods are equilibrated in an optimal way, taking into account the computational cost. In particular, we will prove that the lower convergence rate of the quadrature and finite element method can be achieved up to a logarithmic term as a convergence rate for computing the quantities of interest.

In a further section, we investigate a sparse grid between the truncation of the Karhunen-Loeve expansion and the resolution of the quadrature method. For that, we consider an integration problem over an infinite-dimensional domain. We apply either a Monte Carlo rule or construct a product quadrature which exploits the anisotropy of the integrand and balance it with the truncation to a finite set of variables. This leads to a reduction of the cost when a infinite-dimensional problem is considered.

In the second part of this thesis, we approach the problem from an a-posteriori viewpoint and construct an approximation to the quantities of interest with an adaptive algorithm. We modify the dimension-adaptive combination technique proposed in [7, 8] for product integration and the simulation of polymeric fluids such that the spatial and stochastic approximation is balanced in the computation of quantities of interest for a stochastic partial differential equation. The algorithm is based on a profit indicator. It detects the anisotropy of the stochastic variables and adjusts the quadrature levels. In a further step, the algorithm is extended to problems with an infinite-dimensional stochastic parameter space. The truncation to a finite number of random variables is not assumed to be fixed. Rather the algorithm takes into account the truncation level of the Karhunen-Loeve expansion as an additional direction for the combination technique. In this way, a problem-specific sparse grid is found that includes the truncation, spatial and stochastic approximation.

Finally, the dimension-adaptive combination technique is applied to a problem arising from the study of groundwater flows. This area is of particular interest for the energy sector and management of water. There, the flow through a porous medium is often modeled by a Darcy flow problem. However, as typically only very little data about the sediments is available, the permeability of the sediment is often described with a lognormal random field. Using the dimension-adaptive combination technique, we compute the expectation of the flow and pressure.

The thesis is structured as follows: In Section 2 we will state the problem considered in this thesis more precisely. Subsequently, in Section 3, the numerical methods for the computation of the Bochner integrals, the spatial discretization, and the reduction of a random field to a finite-dimensional stochastic parameter space will be presented. Sections 4 and 5 address the a-priori approach. In Section 4, we will describe the regular sparse grid method and the combination technique. Thereafter, we will consider an anisotropic version of the sparse grid which allows considering methods with different approximation properties in Section 5. The a-posteriori part of this thesis consists of the description of the dimension-adaptive combination technique for fixed and variable number of stochastic variables in Section 6. The thesis concludes with a numerical example for the flow through a porous medium in Section 7.

Detailed references will be given at the beginning of each section. Throughout this thesis, we will avoid the use of generic constants by using the notation  $C \lesssim D$  and  $C \approx D$  if  $C \lesssim D$  and  $D \lesssim C$ .

## 2 Problem Description

As mentioned in the introduction, we will discuss the computation of statistics of the solution of a PDE throughout this thesis. As a model problem we consider the elliptic PDE

$$-\operatorname{div}(a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) = f(\mathbf{x}) \quad \text{in } D \subset \mathbb{R}^n \quad (2)$$

together with zero Dirichlet boundary conditions, whereby the derivatives are only with respect to the spatial variable  $\mathbf{x}$ . Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space, then  $u(\mathbf{x}, \omega)$  is a solution to (2) if the equation (2) is satisfied for almost every  $\omega \in \Omega$ . Throughout the thesis, we denote the spatial domain with  $D \subset \mathbb{R}^n, n = 1, 2$ . If not specified otherwise, we will assume  $D = [0, 1]^n$ .

We consider two cases: either we assume that the diffusion coefficient  $a(\mathbf{x}, \omega)$  depends on a finite set of random variables, or we consider a random field for  $a(\mathbf{x}, \omega)$  and need to reduce it to a finite-dimensional problem.

### 2.1 Finite-dimensional Stochastic

In the case of a finite-dimensional stochastic, we assume that the stochastic dependence is given by a finite number of random variables. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space and let  $\mathbf{y}(\omega) = (y_1(\omega), \dots, y_m(\omega)) \in \mathbb{R}^m$  be a  $m$ -dimensional random vector whose components are independently distributed. Then the diffusion coefficient can be written as

$$a(\mathbf{x}, \omega) = a(\mathbf{x}, \mathbf{y}(\omega)).$$

Instead of considering the PDE (2) for  $\omega \in \Omega$  in the abstract probability space, it is easier to formulate the equation in terms of the variables  $\mathbf{y}$ . Thus, we define the measure  $\mathbb{P}_{\mathbf{y}} = \mathbf{y}_{\#} \mathbb{P}$  as the pushforward of the probability measure on  $\Omega$  through the map

$$\mathbf{y} : \Omega \rightarrow \mathbf{\Gamma}, \quad \omega \mapsto (y_1(\omega), \dots, y_m(\omega))$$

with  $\mathbf{\Gamma} = \Gamma_1 \times \dots \times \Gamma_m$  where  $\Gamma_i$  is the image of  $y_i$  for  $i = 1, \dots, m$ . The diffusion coefficient  $a(\mathbf{x}, \mathbf{y})$  and the solution  $u(\mathbf{x}, \mathbf{y})$  are thereby considered in dependence of  $\mathbf{y}$ , where  $\mathbf{y}$  lies in  $\mathbf{\Gamma}$ , a subset of  $\mathbb{R}^m$ . More generally, the variable  $\mathbf{y}$  does not need to originate from a probabilistic setting, but can describe a dependence of the diffusion coefficient on a set of parameters.

In any case, the PDE (2) is studied in the parametrized version: We aim at finding the function  $u(\mathbf{x}, \mathbf{y})$  such that, for ( $\mathbb{P}_{\mathbf{y}}$ -almost) every point  $\mathbf{y} \in \mathbf{\Gamma}$ , it solves

$$\begin{aligned} -\operatorname{div}(a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y})) &= f(\mathbf{x}) & \mathbf{x} \in D, \\ u(\mathbf{x}, \mathbf{y}) &= 0 & \mathbf{x} \in \partial D, \end{aligned} \quad (3)$$

where the differential operators  $\operatorname{div}$  and  $\nabla$  are again with respect to the spatial variable  $\mathbf{x}$ .

For the distribution of  $\mathbf{y}$  we make the following assumptions. We assume that the probability measure on  $\mathbf{\Gamma}$  is absolutely continuous with respect to the Lebesgue measure, i.e.,  $d\mathbb{P}_{\mathbf{y}} = \rho(\mathbf{y})d\mathbf{y}$ . Moreover, we assume that the probability density  $\rho$  is known explicitly and, requiring independence of the variables, it can be written as

$$\rho(\mathbf{y}) = \prod_{i=1}^m \rho_i(y_i).$$

Further, we need to make assumptions to ensure existence and uniqueness of a solution to the PDE. To that end, we assume two constants  $a_{min}$  and  $a_{max}$  exist such that for all  $\mathbf{x} \in D$  and  $\mathbb{P}_{\mathbf{y}}$ -almost every  $\mathbf{y} \in \mathbf{\Gamma}$

$$0 < a_{min} \leq a(\mathbf{x}, \mathbf{y}) \leq a_{max} < \infty.$$

The Lax-Milgram lemma (see, e.g., [9]) then guarantees well-posedness of (3), in the sense that for  $\mathbb{P}_{\mathbf{y}}$ -almost every  $\mathbf{y} \in \mathbf{\Gamma}$  there exists a unique solution  $u(\mathbf{y}) \in H_0^1(D)$  for  $f \in L^2(D)$ . Furthermore, it holds that

$$\|u(\mathbf{y})\|_{H_0^1(D)} \leq \frac{C}{a_{min}} \|f\|_{L^2(D)}$$

with  $C > 0$  being the Poincaré constant. This ensures that the solution  $u$  is in  $L_\rho^2(\mathbf{\Gamma}, H_0^1(D))$ . It is also possible to consider a right hand side in (3) which depends on a set of random variables. In this case, we need to require  $\|f\|_{L^2(D)} \in L_\rho^2(\mathbf{\Gamma})$  for  $u \in L_\rho^2(\mathbf{\Gamma}, H_0^1(D))$ .

As mentioned in the introduction, we intend to compute the first and second moment of the solution to the PDE. With respect to the push forward measure, computing the quantities of interest corresponds to the evaluation of the integrals

$$\mathbb{E}[u^p(\mathbf{x}, \mathbf{y}(\omega))] = \int_{\Omega} u^p(\mathbf{x}, \mathbf{y}(\omega)) d\mathbb{P}(\omega) = \mathbb{E}_{\mathbf{y}}[u^p(\mathbf{x}, \mathbf{y})] = \int_{\mathbf{\Gamma}} u^p(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{y} \quad (4)$$

for  $p = 1, 2$ . Since  $u \in L_\rho^2(\mathbf{\Gamma}, H_0^1(D))$ , these Bochner integrals are well-defined. We will drop the dependence on  $\mathbf{y}$  in  $\mathbb{E}_{\mathbf{y}}$  in the further notation.

## 2.2 Infinite-dimensional Stochastic

In many applications the diffusion coefficient is described by a random field  $a(\mathbf{x}, \omega)$  instead of depending on a finite number of random variables. In this case, we will approximate the diffusion coefficient by a function that depends only on a finite set of random variables. For that, the random field is written in the Karhunen-Loeve expansion which will be discussed in Section 3.3. The expansion is of the form

$$a(\mathbf{x}, \omega) = \mathbb{E}[a(\mathbf{x})] + \sum_{k=1}^{\infty} \sqrt{\lambda_k} \psi_k(\mathbf{x}) y_k(\omega),$$

where  $y_k$  are uncorrelated random variables. By truncating the series after  $m$  terms, the problem can be reduced to one with a finite-dimensional stochastic.

Also, in the infinite-dimensional case, existence and uniqueness of (3) needs to be assured. It is provided by the Lax-Milgram lemma, when the diffusion coefficient is coercive. Therefore, we assume that the random variables are bounded, for example,  $y_k \in \mathcal{U}(-\sqrt{3}, \sqrt{3})$ , and that the mean  $\mathbb{E}[a(\mathbf{x})]$  is such that there exist constants  $a_{min}, a_{max} > 0$  with

$$\mathbb{P} \left[ a_{min} \leq \operatorname{ess\,inf}_{\mathbf{x} \in D} a(\mathbf{x}, \omega) \leq \operatorname{ess\,sup}_{\mathbf{x} \in D} a(\mathbf{x}, \omega) \leq a_{max} \right] = 1.$$

### 3 Numerical Methods

As mentioned in the introduction, several numerical approximations need to be applied which will be discussed in this section in order to compute the quantities of interest. First, we only treat the case where the stochastic input is finite-dimensional. Although other approaches exist, like a stochastic Galerkin method (e.g., [2]), we focus on methods where the random dependence of  $u$  is approximated by evaluating the function on a set of points  $\mathbf{y}_i \in \Gamma$  which are later combined to approximate the integral. This has the advantage of being able to use deterministic code for solving the PDE. We will start with discussing quadrature methods for evaluating the Bochner integral in Section 3.1. For each quadrature point, the solution to the PDE needs to be approximated as well for which we apply a finite element method discretization in space (see Section 3.2). In the last part, we look at an infinite-dimensional stochastic and discuss the Karhunen-Loeve expansion and finite noise truncation which approximates the random field by a finite number of random variables.

#### 3.1 Quadrature Methods

As stated above, in order to calculate the moments of the solution to the PDE, we must approximate the Bochner integral

$$(\mathbf{I}v)(\mathbf{x}) = \int_{\Gamma} v(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{y} \tag{5}$$

over the  $m$ -dimensional domain  $\Gamma$ , where  $v(\mathbf{x}, \mathbf{y}) = u^p(\mathbf{x}, \mathbf{y})$  with  $p = 1, 2$ . For this purpose, we consider quadrature rules of the type

$$(\mathbf{I}v)(\mathbf{x}) \approx (\mathbf{Q}_N v)(\mathbf{x}) = \sum_{i=1}^N w_i v(\mathbf{x}, \mathbf{y}_i)$$

which are based on the evaluation of the integrand at  $N$  quadrature points  $\mathbf{y}_i \in \Gamma$  that are considered with a quadrature weight  $w_i$ .

We will look at various quadrature rules of this type which arise from two different perspectives: On the one hand, we will consider a sampling method where the

quadrature points are sampled from the domain  $\Gamma$  according to the distribution. On the other hand, deterministic methods are considered, where the function is evaluated on a set of collocation points  $\mathbf{y}_i \in \Gamma$ . Interpolating these solutions  $v(\mathbf{x}, \mathbf{y}_i)$  provides an approximation to the integrand. The integral  $(Iv)(\mathbf{x})$  is then approximated by integration of the interpolant.

Different properties can be studied for quadrature rules to measure the quality of the method. Either the degree of polynomial exactness, which is the maximal degree of polynomials integrated exactly by the quadrature formula, or the convergence rate of the error in terms of the number of quadrature points  $N$  can be evaluated. We will focus on the latter. Since the integral is itself a function in  $\mathbf{x}$ , the quadrature error must be measured in a spatial norm. As  $u \in L^2_\rho(\Gamma, H^1_0(D))$ , we have  $\mathbb{E}[u] \in H^1_0(D)$  and  $\mathbb{E}[u^2] \in W^{1,1}_0(D)$ . Therefore, we require an error bound of the form

$$\|(\mathbf{I} - \mathbf{Q}_N)v\|_{\mathcal{X}} \lesssim \varepsilon(N)\|v\|_{\mathcal{Y}(\Gamma, \mathcal{X})},$$

where  $\mathcal{X} = H^1(D)$  for  $p = 1$  and  $\mathcal{X} = W^{1,1}(D)$  for  $p = 2$ . The space  $\mathcal{Y}(\Gamma, \mathcal{X}) \subset L^2_\rho(\Gamma, \mathcal{X})$  is a suitable Bochner space whose choice depends on the quadrature. The rate of convergence is given by the decay rate of  $\varepsilon(N)$ , where we use  $N$ , the number of quadrature points, as a measure for the cost of the quadrature method. This assumes that each evaluation of the integrand can be done in constant time.

In order to apply the same quadrature method with different approximation power, we define a sequence of quadrature rules  $\mathbf{Q}_l = \mathbf{Q}_{N_l}$  where the number of quadrature points is defined by a level parameter  $l$ . We assume  $N_l < N_{l+1}$  and  $\lim_{l \rightarrow \infty} \mathbf{Q}_l(v) = \mathbf{I}(v)$ , so that the sequence of quadratures applied to a function converges to the integral. As we will see later on, it is sometimes favorable to be able to construct a nested sequence of quadrature rules, where the quadrature points of  $\mathbf{Q}_l$   $\{\mathbf{y}_i\}_{i=1}^{N_l}$  are a subset of the nodes  $\{\mathbf{y}_i\}_{i=1}^{N_{l+1}}$  used for  $\mathbf{Q}_{l+1}$ . However, this property is not required. Hence, we also discuss non-nested quadrature rules.

In the next subsections, we will discuss separately the different quadrature rules used for computing the quantities of interest. The approximation of integrals is of fundamental interest in numerical mathematics. Hence, different quadrature methods have been developed with various properties, which can be found in many textbooks (for example, [10, 11]). Here, we consider three different methods. As they have different properties, their application is convenient in different situations. We start with the dimension-independent Monte Carlo method. Then we move to one-dimensional quadrature rules and discuss the trapezoidal and Gaussian quadrature rule. Finally, the univariate methods can be combined to a tensor product quadrature rule which will be addressed in the last part.

### 3.1.1 Monte Carlo Sampling

One possible way to approximate the integral (5) is to apply a Monte Carlo quadrature. It is often used because little knowledge about the integrand is necessary for convergence.

The Monte Carlo quadrature approximates the integral from a probabilistic point of view. It is based on the law of large numbers which states that the statistical average of a set of  $N$  independent and identically distributed (i.i.d.) random variables converges to the expectation. Thus, the Monte Carlo quadrature points  $\{\mathbf{y}_i\}_{i=1}^{N_y}$  are sampled from the set  $\mathbf{\Gamma}$  independently according to the distribution of  $\mathbf{y}$ , which is given by the density  $\rho$ . In addition, each of the samples is considered to be equally important which is expressed by choosing uniform weights  $w_i = \frac{1}{N}$ . Together, this yields the formula for the Monte Carlo quadrature

$$(\mathbf{Q}_N v)(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N v(\mathbf{x}, \mathbf{y}_i).$$

By construction, the Monte Carlo method is unbiased

$$\mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N v(\mathbf{x}, \mathbf{y}_i) \right] = \frac{1}{N} \sum_{i=1}^N \mathbb{E} [v(\mathbf{x}, \mathbf{y}_i)] = \mathbb{E} [v(\mathbf{x}, \mathbf{y})].$$

Furthermore, a sequence of nested quadrature rules can easily be constructed. This only requires sampling additional points  $\{\mathbf{y}_i\}_{i=N_l+1}^{N_{l+1}}$  independently.

Since random samples are used, not the difference  $\|\mathbf{I}v - \mathbf{Q}_N v\|_{\mathcal{X}}$ , but the root mean square error (RMSE) is considered for quantifying the convergence

$$\text{RMSE} = \sqrt{\mathbb{E} [\|(\mathbf{I} - \mathbf{Q}_N)v\|_{\mathcal{X}}^2]}.$$

A straightforward calculation (cf. [12]) provides the following convergence rate of the mean square error for  $v \in L^2_{\rho}(\mathbf{\Gamma}, \mathcal{X})$

$$\mathbb{E} [\|(\mathbf{I} - \mathbf{Q}_N)v\|_{\mathcal{X}}^2] \leq \frac{1}{N} \|v\|_{L^2_{\rho}(\mathbf{\Gamma}, \mathcal{X})}^2.$$

This implies a convergence rate of  $1/2$ , which is slow compared to other one-dimensional quadrature rules. In order to double the accuracy, the number of samples needs to be quadrupled. This leads to a large number of quadrature points being required to achieve a small error.

Nevertheless, the Monte Carlo quadrature has some advantages. The method is easy to implement and does not require any smoothness assumptions, except for the existence of the variance. Furthermore, the rate of convergence is independent of the dimension of the integration domain  $\mathbf{\Gamma}$ . Hence, it can be applied to high-dimensional problems and provides a better rate than a product quadrature. These major advantages are the reasons why Monte Carlo is often chosen as a quadrature method in applications.

In addition, there are several approaches to improve the error. For the standard Monte Carlo quadrature, various methods, known as variance reduction techniques, aim at reducing the constant in the error estimate. These approaches replace the integrand or adjust the sampling such that the variance of the integrand is smaller, while maintaining the same mean value. Among the most

commonly used techniques are methods such as antithetic variables, importance sampling, control variates and stratification. However, with all these techniques, the convergence rate of  $1/2$  is not affected ([13, 2]).

Another approach, the Quasi Monte Carlo method, is based on replacing the random samples by a set of deterministic points. There, a sequence of points is designed to cover the domain  $\Omega$  in a good way, measured by the concept of low discrepancy. Different sequences exist, including Halton, Hammersley and Sobol points. Provided there is some regularity of the integrand, the Quasi Monte Carlo methods achieve an error rate of  $\mathcal{O}(N^{-1} \log(N)^d)$ , depending on the discrepancy of the point set. For more details, see for example, [2] or [14], and for the computation of moments of a PDE [15]. We will not discuss these methods further here. Instead, we will focus next on a different type of quadrature rules.

### 3.1.2 Univariate Quadrature Rules and Product Quadrature

Instead of sampling points in the entire parameter space  $\Gamma$ , the integral (5) can be regarded as an integration problem in each parameter  $y_k$  for  $k = 1, \dots, m$ . An  $m$ -dimensional quadrature rule is then obtained by the combination of univariate quadratures.

Therefore, we discuss in the following the approximation of an integral over a one-dimensional domain first. We look at two commonly used types of one-dimensional quadrature rules which both rely on an interpolation of the integrand. Later, the univariate rules will be combined to a tensor product quadrature rule.

#### Newton Cotes

The class of *Newton Cotes* rules are integration methods for an integration problem over a bounded interval  $\Gamma = [t_1, t_2]$ . They are based on polynomial interpolation of the integrand on equidistant nodes. The corresponding weights are obtained by integration of the Lagrange basis polynomials. Thus, depending on the number of quadrature points  $N$ , the Newton Cotes formulas have the polynomial degree of exactness  $N - 1$ . Furthermore, Newton Cotes formulas can be constructed such that for a function  $v \in C^r(\Gamma, \mathcal{X})$ ,  $r \in \mathbb{N}$

$$\|(I - Q_N)v\|_{\mathcal{X}} \lesssim N^{-r} \|v\|_{C^r(\Gamma, \mathcal{X})}.$$

Examples for Newton Cotes formulas are the trapezoidal and Simpson rules. However, since the interpolation polynomials tend to have large oscillations when the number of nodes is large, low order composite rules are often applied, which use piecewise polynomial interpolations of the integrands instead. Although various composite Newton Cotes rules could be investigated, we will consider only the composite trapezoidal rule, which takes the form

$$(Q_N v)(\mathbf{x}) = \sum_{i=0}^N w_i v(\mathbf{x}, y_i)$$

with  $y_i = t_1 + ih$ , where  $h = \frac{t_2 - t_1}{N}$  and  $w_i = \frac{h}{2}$  for  $i = 0$ , and  $i = N$  and  $w_i = h$  otherwise.

For a function  $v \in C^2(\Gamma, \mathcal{X})$ , the composite trapezoidal rule satisfies the error bound

$$\|(I - Q_N)v\|_{\mathcal{X}} \lesssim N^{-2} \max_{y \in \Gamma} \|\partial_y^2 v(y)\|_{\mathcal{X}}. \quad (6)$$

Furthermore, since equidistant quadrature points are applied, a nested sequence of trapezoidal rules can be defined. Choosing  $N_l = 2^l$ , it holds that  $\{y_i\}_{i=0}^{N_l} \subset \{y_i\}_{i=0}^{N_{l+1}}$  and the quadrature error of  $Q_l$  is of order  $\mathcal{O}(2^{-2l})$ .

## Gaussian Quadrature

As a second class of univariate quadrature rules, Gaussian quadrature methods will be considered. These rules are also based on an interpolation among a set of points, which, however, is chosen such that the highest degree of polynomial exactness can be achieved. In this way, a quadrature rule with  $N$  nodes can be constructed that is exact for polynomials up to degree  $2N - 1$ .

Different to the Newton Cotes formulas, we do not require a bounded set  $\Gamma$  and the integration problem can be generalized to a weighted integral

$$\int_{\Gamma} v(y) \rho(y) dy.$$

For the approximation of this integral, the function  $v$  is replaced by a global polynomial that interpolates  $v$  at a set of quadrature points  $\{y_i\}_{i=1}^N$  and the weights are computed according to

$$w_i = \int_a^b \ell_i^{\{y_j\}_j}(t) \rho(t) dt,$$

where  $\ell_i^{\{y_j\}_j}(t)$  are the Lagrangian interpolation polynomials.

To achieve the highest possible degree of polynomial exactness, the  $N$  roots of the polynomial  $p_N$  are chosen as quadrature nodes, where  $p_N$  is the orthogonal polynomial of order  $N$  with respect to the inner product

$$(v, w)_{L_{\rho}^2(\Gamma)} = \int_{\Gamma} v(y) w(y) \rho(y) dy$$

in the space  $L_{\rho}^2(\Gamma)$ . Numerically, the quadrature nodes and weights can be computed using a three-term recursion of the orthogonal polynomials and solving an eigenvalue problem (see, for example, [11]).

The rules are based on different orthogonal polynomials, depending on  $\rho$ . Two cases are relevant for us: Throughout the thesis we mainly assume  $y$  to be a uniform random variable with values in the bounded interval  $\Gamma = [t_1, t_2]$ , for which the density is given by  $\rho \equiv \frac{1}{t_2 - t_1}$ . Using an affine transformation, the integral can be reformulated as an integral over  $[-1, 1]$  with a constant weight function  $\rho$ , for which the corresponding orthogonal polynomials are the Legendre polynomials. For this reason, the Gaussian quadrature in this case is called Gauss-Legendre rule. For a random variable following the standard normal distribution, the probability density function is given by  $\rho(y) = C \exp(-\frac{y^2}{2})$  and  $\Gamma = (-\infty, \infty)$ . Hence, the Hermite polynomials are used and the resulting quadrature is the Gauss-Hermite quadrature.

Both the Gauss-Legendre and Gauss-Hermite quadrature nodes are non-nested. There are nested versions, like the Clenshaw-Curtis rule [16] and extensions of Gaussian rules [17, 18], but they are not considered here, as nestedness is not required.

Once the quadrature weights and nodes are provided, the Bochner integral

$$\int_{\Gamma} v(\mathbf{x}, y) \rho(y) dy$$

is approximated by

$$(Q_N v)(\mathbf{x}) = \sum_{i=1}^N w_i v(\mathbf{x}, y_i).$$

In the case of the Gauss-Legendre quadrature, we consider the quadrature as an operator  $Q_N : C^0(\Gamma, \mathcal{X}) \rightarrow \mathcal{X}$  with  $C^0(\Gamma, \mathcal{X}) \subset L^2_{\rho}(\Gamma, \mathcal{X})$ . The operator is continuous with constant 1 ([19]). Furthermore, the following convergence rate stated in [20, 21] holds true, which is based on the approximation analysis of the interpolating function in  $L^2_{\rho}(\Gamma, \mathcal{X})$  ([19]).

**Lemma 1.** *If the function  $v : \Gamma \rightarrow \mathbb{R}$  is analytically extendable into a region in the complex plane  $\Sigma(\Gamma, \tau) = \{z \in \mathbb{C} : \text{dist}(z, \Gamma) \leq \tau\}$  the Gauss-Legendre quadrature error satisfies*

$$\|(I - Q_N)v\|_{\mathcal{X}} \leq g(\tau) \exp(-h(\tau)(2N - 1)) \|v\|_{C^0(\Sigma(\Gamma, \tau), \mathcal{X})}$$

where  $g(\tau) = \frac{4}{\kappa - 1}$  and  $h(\tau) = \log(\kappa)$  with  $\kappa = \frac{2\tau}{|\Gamma|} + \sqrt{1 + \left(\frac{2\tau}{|\Gamma|}\right)^2}$ .

Similar results of exponential convergence hold for other Gaussian quadratures as well, but they require the analysis in a different weighted function space  $C^0_{\sigma}(\Gamma, \mathcal{X}) \subset L^2_{\rho}(\Gamma, \mathcal{X})$  (see [19] for details).

Lemma 1 implies an exponential convergence rate of the Gaussian quadrature for smooth functions. If instead the integrand is of limited regularity, the convergence rate reduces to an algebraic rate (see, e.g., [22]).

Before concluding the section with discussing the product quadrature, we provide a condition that ensures that the assumption in Lemma 1 is satisfied. As we later want to combine the univariate Gaussian quadrature rules to a product rule,  $v(\mathbf{y})$  needs to be extendable for all  $y_1, \dots, y_m$ . For that, we introduce the notation

$$\mathbf{y}_k^* = (y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_m) \in \mathbb{R}^{m-1}$$

and

$$\Gamma_k^* = \bigtimes_{\substack{i=1 \\ i \neq k}}^m \Gamma_i. \quad (7)$$

The complex extendibility can be ensured by the following Lemma (cf. Lemma 6.3 in [23]).

**Lemma 2.** *Let  $v \in C^0(\Gamma)$  and let the derivatives of  $v$  satisfy*

$$\|\partial_{y_k}^j v\|_{C^0(\Gamma)} \lesssim j! \mu_k^j$$

*with some  $\mu_k > 0$ . Then, for  $\tau_k \in (0, \frac{1}{\mu_k})$ , the function*

$$v : \Gamma_k \rightarrow \mathbb{R}, y_k \mapsto v(y_k, \mathbf{y}_k^*)$$

*admits an analytic extension into  $\Sigma(\Gamma_k, \tau_k) = \{z \in \mathbb{C} : \text{dist}(z, \Gamma_k) \leq \tau_k\}$ . The function  $v$  is bounded with respect to the norm*

$$\|v\|_{C^0(\Sigma(\Gamma_k, \tau_k); C^0(\Gamma_k^*))} = \sup_{z \in \Sigma(\Gamma_k, \tau_k)} \|v(z)\|_{C^0(\Gamma_k^*)}.$$

The proof follows by the same arguments as in [23].

## Tensor Product Quadrature

A straightforward way to obtain a quadrature rule over an  $m$ -dimensional domain is the tensorization of univariate quadrature rules. For  $m$  univariate quadratures  $Q_{l_i}^{(i)}$  with  $N_{l_i}$ ,  $i = 1, \dots, m$ , quadrature nodes, we define the product quadrature as

$$\begin{aligned} (\mathbf{Q}_l v)(\mathbf{x}) &= \left( Q_{l_1}^{(1)} \otimes \dots \otimes Q_{l_m}^{(m)} v \right) (\mathbf{x}) \\ &= \sum_{i_1=1}^{N_{l_1}} \dots \sum_{i_m=1}^{N_{l_m}} w_{i_1}^{(l_1)} \dots w_{i_m}^{(l_m)} v \left( \mathbf{x}, y_{i_1}^{(l_1)}, \dots, y_{i_m}^{(l_m)} \right) \end{aligned}$$

with  $\mathbf{l} = (l_1, \dots, l_m)$ . Hence, the number of quadrature points for  $\mathbf{Q}_l$  is  $\prod_{i=1}^m N_{l_i}$ .

The error can be bounded by the sum of the errors of the one-dimensional quadrature rules using the continuity of the quadrature and integral operators.

$$\begin{aligned}
& \|(\mathbf{I} - \mathbf{Q}_l)v\|_{\mathcal{X}} \\
& \leq \sum_{k=1}^m \left\| \left( Q_{l_1}^{(1)} \otimes \dots \otimes Q_{l_{k-1}}^{(k-1)} \otimes \left( I^{(k)} - Q_{l_k}^{(k)} \right) \otimes I^{(k+1)} \otimes \dots \otimes I^{(m)} \right) v \right\|_{\mathcal{X}} \\
& \lesssim \sum_{k=1}^m \sup_{\mathbf{y}_k^* \in \Gamma_k^*} \left\| \left( I^{(k)} - Q_{l_k}^{(k)} \right) v(\mathbf{y}_k^*) \right\|_{\mathcal{Y}(\Gamma_k, \mathcal{X})}, \tag{8}
\end{aligned}$$

where the norm  $\|\cdot\|_{\mathcal{Y}(\Gamma_k, \mathcal{X})}$  depends on the one-dimensional quadrature rules. We will discuss in detail the construction of an anisotropic product quadrature consisting of one-dimensional Gauss-Legendre rules in Section 5.2.2.

In view of (8), the errors of all univariate quadrature rules should be balanced since the largest will dominate the overall error of the product quadrature. This, however, leads to computationally expensive methods as the dimensions increase. To illustrate this, we consider a univariate quadrature rule with  $N$  points that achieves an error of  $\varepsilon = N^{-r}$ . Applying the same rule in all  $m$  directions leads to a product quadrature with an error of  $\varepsilon$ . However,  $\mathbf{N} = N^m$  quadrature points have been used for achieving this error. The convergence rate in terms of computational cost  $\mathbf{N}$ , hence, reduces for the product quadrature to

$$\|(\mathbf{I} - \mathbf{Q})v\|_{\mathcal{X}} = \mathcal{O}(\mathbf{N}^{-r/m}).$$

This phenomenon, that for the same accuracy the number of points grows exponentially with the dimensionality of the problem, is referred to as the *curse of dimensionality* [24]. The same problem occurs for other approximations that rely on a tensor product of numerical approximations.

After having discussed the approximation in the parameters  $\mathbf{y}$  for the computation of the integrals, we will address the spatial discretization in the next section.

### 3.2 Finite Element Method

For each quadrature or sample point  $\mathbf{y} \in \mathbf{\Gamma}$  the function  $u(\mathbf{y}) \in H_0^1(D)$  must be evaluated. Since  $u$  is the solution to an elliptic PDE, this means solving a deterministic PDE for each sample point. In general, this can only be done approximately. Here, we consider the case where the solution  $u(\mathbf{y})$  is approximated by a linear finite element method. Other finite elements or a spectral method, instead of a local approach, could also be considered.

For the finite element method, we consider the PDE (3) in the weak formulation where a weak solution satisfies

$$\int_D a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y}) \nabla v(\mathbf{x}) d\mathbf{x} = \int_D f(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}$$

for all  $v \in H_0^1(D)$ . An approximation to the solution can be obtained by considering a finite-dimensional subspace  $\mathcal{V}_h \subset H_0^1(D)$ . We define for a uniform and regular triangulation  $\mathcal{T}_h$  the space of piecewise linear functions

$$\mathcal{V}_h = \{v \in C(D) : v|_{\partial D} = 0 \text{ and } v|_K \in \mathcal{P}_1 \ \forall K \in \mathcal{T}_h\},$$

where  $\mathcal{P}_1$  denotes the space of linear functions and find  $u_h(\mathbf{y}) \in \mathcal{V}_h$  such that

$$\int_D a(\mathbf{x}, \mathbf{y}) \nabla u_h(\mathbf{x}, \mathbf{y}) \nabla v_h(\mathbf{x}) d\mathbf{x} = \int_D f(\mathbf{x}) v_h(\mathbf{x}) d\mathbf{x}$$

for all  $v_h \in \mathcal{V}_h$ .

From standard theory for piecewise linear finite elements, the approximation properties of the finite element solution  $u_h(\mathbf{y})$  are known (see, e.g., [25, 26, 27]).

**Lemma 3.** *Let  $h = \max_{K \in \mathcal{T}_h}(\text{diam}(K))$ . If  $u \in C^0(\bar{D}) \cap H^{s+1}(D)$  with  $s > 0$ , then*

$$\begin{aligned} \|u(\mathbf{y}) - u_h(\mathbf{y})\|_{H^1(D)} &\leq Ch^{\min\{1,s\}} \|u(\mathbf{y})\|_{H^{\min\{2,s+1\}}(D)}, \\ \|u(\mathbf{y}) - u_h(\mathbf{y})\|_{L^2(D)} &\leq Ch^{\min\{2,s+1\}} \|u(\mathbf{y})\|_{H^{\min\{2,s+1\}}(D)}. \end{aligned}$$

Since for  $f \in L^2(D)$  the PDE solution  $u(\mathbf{y})$  is in  $H^2(D)$ , Lemma 3 implies the following statement.

**Lemma 4.** *Let  $h = \max_{K \in \mathcal{T}_h}(\text{diam}(K))$ . If  $f \in L^2(D)$ , the finite element approximation satisfies*

$$\|u(\mathbf{y}) - u_h(\mathbf{y})\|_{H^1(D)} \lesssim h \|f\|_{L^2(D)}$$

and

$$\|u^2(\mathbf{y}) - u_h^2(\mathbf{y})\|_{W^{1,1}(D)} \lesssim h \|f\|_{L^2(D)}^2.$$

We will need to apply a sequence of finite element solutions for the sparse grid approach discussed in Section 4. Hence, we consider for the rectangular domain  $D$  a sequence of uniform meshes  $\{\mathcal{T}_l\}_l$  with mesh widths  $h_l = 2^{-l}$  and consider the sequence of piecewise finite element spaces based on these meshes. Then, the convergence rate for  $u_h$  in the  $H^1(D)$ -norm and  $u_h^2$  in the  $W^{1,1}(D)$ -norm is  $\mathcal{O}(2^{-l})$ .

For the numerical examples in this thesis, we will use the finite element software FEniCS ([28], version 2019.1.0) which is designed to solve partial differential equations efficiently. As the collocation and sampling methods combine the evaluations of  $u$  at a set of points  $\{\mathbf{y}_i\}_i$ , we can use the available deterministic code for solving the PDE.

### 3.3 Finite Noise Truncation

So far, we have considered the numerical methods for a problem where the stochastic dependence is given by a finite number of parameters. In the case where the stochastic dependence is infinite-dimensional, we need to approximate in addition the stochastic input by a finite-dimensional object. Therefore, we discuss in this section how to approximate the random field  $a(x, \omega)$  by a finite

number of random variables. For that, we consider a truncation of the Karhunen-Loeve expansion, but mention that there are also other expansions, including more general polynomial chaos expansions and expansions using wavelets (cf. [2]) that could be considered for the approximation of the random field.

The Karhunen-Loeve expansion separates the stochastic and deterministic dependency of the random field and can be seen as a basis representation of  $a$  in  $L^2_{\mathbb{P}}(\Omega) \otimes L^2(D) \cong L^2_{\mathbb{P}}(\Omega, L^2(D))$ . Furthermore, it is related to the singular value decomposition for matrices. The random field is written in the form

$$a(\mathbf{x}, \omega) = \sum_{k \in \mathbb{N}} \psi_k(\mathbf{x}) Z_k(\omega) \quad (9)$$

where  $Z_k$  are random variables and  $\psi_k$  are suitable functions on  $D$ . The Karhunen-Loeve expansion is particularly nice, as it decomposes  $a$  in a bi-orthogonal way, such that the functions  $\psi_k$  are orthogonal in  $L^2(D)$ , and also the  $Z_n$  are orthogonal in the probability space ([2]).

For the description of the Karhunen-Loeve expansion, we follow parts of [4] and [2]. We assume that the random field is included in the Bochner space  $L^2_{\mathbb{P}}(\Omega; L^2(D))$ . Hence, the expectation

$$\mathbb{E}[a(\mathbf{x})] = \int_{\Omega} a(\mathbf{x}, \omega) d\mathbb{P}(\omega)$$

and covariance function

$$\text{cov}_a(\mathbf{x}, \mathbf{x}') = \int_{\Omega} (a(\mathbf{x}, \omega) - \mathbb{E}[a(\mathbf{x})]) (a(\mathbf{x}', \omega) - \mathbb{E}[a(\mathbf{x}')]) d\mathbb{P}(\omega)$$

are well-defined. We further assume that the expectation and covariance functions are known and that the covariance function is continuous, i.e.,  $\text{cov}_a \in C(D \times D)$ .

We define a linear operator  $C : L^2(D) \rightarrow L^2(D)$  by

$$(Cv)(\mathbf{x}) = \int_D \text{cov}_a(\mathbf{x}, \mathbf{x}') v(\mathbf{x}') d\mathbf{x}',$$

which is a symmetric and positive semi-definite Hilbert-Schmidt operator.

Then, as a consequence of Mercer's theorem (see, e.g., [2]), we have the following statement.

**Theorem 1.** *Under the above assumptions, there exist a sequence of values  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq \dots \geq 0$  with  $\lim_{k \rightarrow \infty} \lambda_k = 0$  and functions  $\psi_k : D \rightarrow \mathbb{R}$  such that*

$$\int_D \text{cov}_a(\mathbf{x}, \mathbf{x}') \psi_k(\mathbf{x}') d\mathbf{x}' = \lambda_k \psi_k(\mathbf{x}) \quad (10)$$

and

$$\int_D \psi_i(\mathbf{x}) \psi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij}.$$

Furthermore, the random field  $a(\mathbf{x}, \omega)$  can be written as an infinite series

$$a(\mathbf{x}, \omega) = \mathbb{E}[a(\mathbf{x})] + \sum_{i=k}^{\infty} \sqrt{\lambda_k} \psi_k(\mathbf{x}) y_k(\omega) \quad (11)$$

where the sum converges in the space  $L^2_{\mathbb{P}}(\Omega, L^2(D))$ .

The random variables  $y_k$  are uncorrelated, centered random variables with unit variance. They are defined by

$$y_k(\omega) = \frac{1}{\sqrt{\lambda_k}} \int_D (a(\mathbf{x}, \omega) - \mathbb{E}[a(\mathbf{x})]) \psi_k(\mathbf{x}) d\mathbf{x}. \quad (12)$$

In other words,  $\psi_k(\mathbf{x})$  and  $\lambda_k$  are the eigenfunctions and eigenvalues of the covariance operator  $C$ , respectively, and  $\psi_k(\mathbf{x})$  and  $y_k(\omega)$  form orthonormal sequences in  $L^2(D)$  and  $L^2(\Omega)$ . Another consequence of Mercer's theorem is uniform convergence of the series (11).

**Theorem 2.** *The series expansion (11) converges uniformly in  $\mathbf{x}$ , i.e.,*

$$\sup_{\mathbf{x} \in D} \mathbb{E} \left[ \left( a(\mathbf{x}, \omega) - \sum_{k=1}^m \sqrt{\lambda_k} \psi_k(\mathbf{x}) y_k(\omega) \right)^2 \right] = \sup_{\mathbf{x} \in D} \sum_{k=m+1}^{\infty} \lambda_k \psi_k(\mathbf{x})^2 \rightarrow 0$$

as  $m \rightarrow \infty$ .

As the distribution of the random diffusion coefficient is often not explicitly known, the equation (12) cannot always be used. Instead, it is often assumed that the random variables are independently distributed. In the case of a Gaussian field, this assumption is automatically satisfied, as uncorrelated normal variables are also independent, but random variables are unbounded. In order to ensure the coercivity for the existence of solutions to the PDE (cf. Section 2.2), we assume that  $y_k$  is uniformly distributed in  $(-\sqrt{3}, \sqrt{3})$ , such that  $\mathbb{E}[y_k] = 0$  and  $\mathbb{E}[y_k^2] = 1$ .

In the course of this thesis, we will make use of the size of the coefficient of  $y_k$ . Hence, we define, in addition, the sequence

$$\gamma_k := \sqrt{\lambda_k} \|\psi_k(\mathbf{x})\|_{L^\infty(D)} \quad (13)$$

and assume that it is summable. The coefficients  $\gamma_k$  then describe an anisotropy with respect to the random variables  $y_k$ . Since the  $y_k$  are identically distributed, the variability of the field  $a$  due to parameter  $y_k$  is small, if the coefficient is small. Hence, the influence of the random variables is not equally strong.

## Truncation

In practice, to avoid infinite-dimensional stochastic domains, the Karhunen-Loeve expansion is truncated after  $m$  terms, for which we write

$$a_m(\mathbf{x}, \omega) = \mathbb{E}[a(\mathbf{x})] + \sum_{k=1}^m \sqrt{\lambda_k} \psi_k(\mathbf{x}) y_k(\omega).$$

By this truncation, the random dependence is reduced to a finite number of random variables. Hence, we are in the setting as for the finite-dimensional stochastic and the randomness can be parameterized, like in Section 2.1.

However, this truncation introduces an error. Measured in the  $L^2_{\mathbb{P}}(\Omega, L^2(D))$ -norm, the error of the  $m$ -term truncation depends on the eigenvalues. Due to the orthonormality of  $\{\psi_k\}$  and  $\{y_k\}$ , it holds

$$\begin{aligned} \|a(\mathbf{x}, \omega) - a_m(\mathbf{x}, \omega)\|_{L^2_{\mathbb{P}}(\Omega, L^2(D))}^2 &= \left\| \sum_{i=m+1}^{\infty} \sqrt{\lambda_i} \psi_i(\mathbf{x}) y_i(\omega) \right\|_{L^2_{\mathbb{P}}(\Omega, L^2(D))}^2 \\ &= \sum_{i,j=m+1}^{\infty} \int_{\Omega} \int_D \sqrt{\lambda_j} \sqrt{\lambda_i} \psi_i \psi_j y_i y_j \, d\mathbf{x} \, d\mathbb{P}(\omega) = \sum_{i=m+1}^{\infty} \lambda_i. \end{aligned}$$

Hence, the decay of the eigenvalues controls the truncation error. In addition, the Karhunen-Loeve expansion satisfies an optimality property in  $L^2_{\mathbb{P}}(\Omega, L^2(D))$ . It is optimal under all possible decompositions of the random field of the form (9):

$$\{y_i, \psi_i\}_{i=1}^m = \underset{\substack{(\xi_i, \phi_i) \\ \int_D \phi_i \phi_j = \delta_{ij}}}{\operatorname{argmin}} \mathbb{E} \left[ \int_D \left( a(\mathbf{x}, \omega) - \mathbb{E}[a(\mathbf{x})] - \sum_{i=1}^m \xi_i(\omega) \phi_i(\mathbf{x}) \right)^2 \, d\mathbf{x} \right].$$

However, if we consider other norms, there is no such optimality result. In fact, other expansions may be better suited. For example, in [29], the authors proved the non-optimality of the Karhunen-Loeve representation for the solution of a PDE with lognormal diffusion coefficient in a concrete example. Instead, they found that the Levy-Ciesielski representation, where  $\psi_k$  are only locally supported and are not orthogonal in  $L^2(D)$ , yields better convergence rates for the approximation of  $u(\mathbf{y})$ . Despite the fact that the Karhunen-Loeve expansion might not be optimal, we will use this expansion for the approximation of random fields in this thesis.

As our goal is to compute the moments of the PDE solution, the truncated version will eventually be used instead of the random field in (3). The error induced by this has been investigated, for example, in [30] for the solution  $u$  of the PDE, where the following was proven:

**Theorem 3.** *The solution  $u_m$  to (3) with the random field  $a_m$  given by the Karhunen-Loeve expansion truncated after  $m$  terms satisfies*

$$\|u(\mathbf{y}) - u_m(\mathbf{y})\|_{H^1(D)} \lesssim \|f\|_{L^2(D)} \sum_{k=m+1}^{\infty} \gamma_k$$

with  $\gamma_k$  as in (13) and a constant which depends on  $a_{\min}$ , but is independent of  $m$ .

Hence, the decay properties of  $\lambda_k$  and  $\gamma_k$  determine the error made by the truncation. Furthermore, these properties affect the stochastic dimension of the

parametric PDE (3). If the coefficients decay fast, less terms are needed in the Karhunen-Loeve representation to obtain a small error. In contrast, if the eigenvalues decay slowly, more terms are needed, resulting in a high-dimensional parametrized problem.

## Decay of Eigenvalues

Next, we discuss the decay of the coefficients  $\lambda_k$  and  $\gamma_k$ . In general, the decay rate for  $\{\lambda_k\}$  depends on the smoothness of the covariance function (see [31, 32, 33], where the authors provided algebraic convergence rates for  $a \in L^2_{\mathbb{P}}(\Omega, H^s(D))$ ). Here, we will examine the decay rates for the covariance functions used in our numerical examples. We consider a certain class of random fields, where the covariance is described by the class of Matérn covariance functions. These functions are often considered, as the regularity and eigenvalue decay is known. The Matérn covariance kernels are defined by a stationary covariance function (see, e.g., [34]).

**Definition 1.** Let  $r = |\mathbf{x} - \mathbf{x}'|_2$  denote the distance between two points and let  $\rho, \sigma^2 > 0$  be the correlation length and variance, respectively. Then the Matérn covariance function of order  $\nu > 0$  is defined as

$$\text{cov}(r; \nu) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \sqrt{2\nu} \frac{r}{\rho} \right)^\nu K_\nu \left( \sqrt{2\nu} \frac{r}{\rho} \right), \quad (14)$$

where  $\Gamma$  is the gamma function and  $K_\nu$  the modified Bessel function of the second kind of order  $\nu$ .

The formula (14) simplifies for  $\nu = s + \frac{1}{2}$  with  $s \in \mathbb{N}$  to

$$\text{cov} \left( r; s + \frac{1}{2} \right) = \sigma^2 \exp \left( -\frac{\sqrt{2s+1}r}{\rho} \right) \frac{s!}{(2s)!} \sum_{i=0}^s \frac{(s+i)!}{i!(s-i)!} \left( \frac{2\sqrt{2s+1}r}{\rho} \right)^{s-i}.$$

Furthermore, in the limit  $\nu \rightarrow \infty$ , the function converges to a Gaussian covariance function

$$\text{cov}(r) = \sigma^2 \exp \left( -\frac{1}{2} \frac{r^2}{\rho^2} \right).$$

The parameter  $\nu$  can be regarded as a smoothness parameter, as it controls the smoothness of the covariance function at the point  $r = 0$ .

For this class of covariance functions, the eigenvalues of (10) satisfy (see [33])

$$\lambda_k \lesssim k^{-(1+2\nu/n)} \quad \forall k \geq 1,$$

where  $n$  is the dimension of  $D$ .

Furthermore, from [15] we have the decay rate for  $\{\gamma_k\}$ . For  $\nu > n$ , the sequence  $\{\gamma_k\}$  is summable and for all  $\varepsilon > 0$

$$\gamma_k \lesssim k^{-\frac{\nu}{n} + \varepsilon}.$$

Hence, the dimension of the domain and smoothness of the covariance function determine the asymptotic decay rates of the sequences, whereby smoother covariance functions lead to a faster decay. The pre-asymptotic behavior is influenced by the correlation length  $\rho$ . Shorter correlation lengths increase the pre-asymptotic domain.

For the limiting case, the Gaussian covariance function is analytic and, as a consequence, exponential decay rates for the eigenvalues hold (see [32] Theorem 2.19)

$$\lambda_k \leq c_1 e^{-c_2 k} \quad \forall k \geq 1$$

with constants depending on the variance and correlation length, whereby a longer correlation length implies a faster decrease. In addition, the eigenfunctions are also analytic, and uniformly bounded in  $L^\infty(D)$  such that the sequence  $\{\gamma_k\}$  also decays exponentially.

### Computation of Eigenpair

For most covariance functions, there is no analytical solution to the eigenvalue problem available. Only in special cases are exact solutions known. In general, however, the Karhunen-Loeve expansion can only be determined numerically. Here, we briefly discuss the approximation that was used for most of the numerical examples based on the work in [35, 4].

The eigenvalue problem (10) is solved using a Galerkin approximation. Instead of looking for eigenfunctions on the infinite-dimensional space  $L^2(D)$ , the search is restricted to a finite-dimensional subset  $V_h \subset L^2(D)$ . The projection onto  $V_h$  we denote by  $P_{V_h} : L^2(D) \rightarrow V_h$ . and assume further that  $\{\phi_1, \dots, \phi_K\}$  is a basis of  $V_h$ . In order to find an approximative eigenpair  $(\lambda_j^h, \psi_j^h)$ , we solve

$$\langle C\psi_j^h - \lambda_j^h \psi_j^h, \phi_i \rangle_{L^2(D)} = 0 \quad \forall i = 1, \dots, K \quad (15)$$

which is

$$\int_D \left( \int_D \text{cov}(\mathbf{x}, \mathbf{x}') \psi_j^h(\mathbf{x}') d\mathbf{x}' \right) \phi_i(\mathbf{x}) d\mathbf{x} = \lambda_j^h \int_D \psi_j^h(\mathbf{x}) \phi_i(\mathbf{x}) d\mathbf{x} \quad \forall i = 1, \dots, K.$$

Since  $\psi_j^h \in V_h$ , it can be written as  $\psi_j^h = \sum_{k=1}^n \alpha_{j,k} \phi_k$ . Plugging this representation into (15), we need to solve the matrix eigenvalue problem

$$\mathbf{W}\boldsymbol{\alpha}_j = \lambda_j^h \mathbf{M}\boldsymbol{\alpha}_j$$

where

$$W_{ij} = \int_D \int_D \text{cov}(\mathbf{x}, \mathbf{x}') \phi_i(\mathbf{x}) \phi_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}' \quad \text{and} \quad M_{ij} = \int_D \phi_i(\mathbf{x}) \phi_j(\mathbf{x}) d\mathbf{x}$$

with  $\boldsymbol{\alpha}_j = (\alpha_{j,1}, \dots, \alpha_{j,K})^T$ . The matrix  $\mathbf{W}$  is symmetric, semi-positive definite and  $\mathbf{M}$  is symmetric and positive definite. For this type of generalized eigenvalue

problem, there are standard iterative solvers available which compute the largest eigenvalues.

However, the entries of  $\mathbf{W}$  are usually not accessible and need to be approximated. For that, we choose for  $V_h$  the finite element space with piecewise linear functions and approximate the covariance function by

$$\text{cov}(\mathbf{x}, \mathbf{x}') \approx \text{cov}_h(\mathbf{x}, \mathbf{x}') = \sum_{i,j=1}^n C_{i,j} \phi_i(\mathbf{x}) \phi_j(\mathbf{x}'),$$

where  $C_{ij}$  are the entries of matrix  $\mathbf{C}$  with  $C_{ij} = \text{cov}(\mathbf{x}_i, \mathbf{x}_j)$ . The matrix  $\mathbf{W}$  is then approximated by

$$\begin{aligned} W_{ij} &\approx \int_D \int_D \text{cov}_h(\mathbf{x}, \mathbf{x}') \phi_i(\mathbf{x}) \phi_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}' \\ &= \sum_{k,p=1}^n \int_D \int_D \phi_k(\mathbf{x}) \phi_i(\mathbf{x}) C_{k,p} \phi_p(\mathbf{x}') \phi_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}' \\ &= (\mathbf{MCM})_{ij}. \end{aligned}$$

The mass matrix  $\mathbf{M}$  is available in standard finite element software such that the the generalized eigenvalue problem

$$\mathbf{MCM}\boldsymbol{\alpha}_j = \lambda_j^h \mathbf{M}\boldsymbol{\alpha}_j$$

can be solved. Although solving the eigenvalue problem (10) only approximately is an additional approximation to the truncation of the Karhunen-Loeve expansion, we do not pay attention to this error. We use a fine discretization of the space for  $V_h$  and we assume that the eigenvalues and eigenfunctions are close enough to the exact values.

## 4 Sparse Grids and the Combination Technique

After having discussed the different numerical methods separately, we now look at the combination of these methods. As our goal is to compute the first and second moment of the solution to the PDE as quantities of interest, an approximation to that can be obtained by applying successively a truncation of the Karhunen-Loeve expansion for the diffusion coefficient, a finite element approximation of the PDE solution, and a quadrature method. This can be written as

$$\mathbb{E}[u^p] \approx \mathbf{Q}u_{m,h}^p, \quad (16)$$

where  $\mathbf{Q}$  is a quadrature method and  $u_{m,h}$  is the finite element solution to the PDE with the random field truncated to  $m$  terms.

The error of this computation can be split into different parts

$$\|\mathbb{E}[u^p] - \mathbf{Q}u_{m,h}\| \leq \|\mathbb{E}[u^p] - \mathbb{E}[u_m^p]\| + \|\mathbb{E}[u_m^p] - \mathbb{E}[u_{m,h}^p]\| + \|\mathbb{E}[u_{m,h}^p] - \mathbf{Q}u_{m,h}^p\|$$

where  $u_m$  is the solution of the PDE with a truncated random field  $a_m$ , while  $u_{m,h}$  and  $u_h$  denote the finite element approximation of  $u_m$  and  $u$ , respectively. The first term in the error bound is due to a truncation of the Karhunen-Loeve expansion, the second term is the error caused by the finite element discretization, and the third term describes the quadrature error.

In the case where the diffusion coefficient of the PDE depends only on finitely many random variables, the first error source does not occur. Furthermore, beyond the terms in (16), the quadrature error itself might consist of several terms. If a product of one-dimensional quadrature rules is used for the  $m$ -dimensional integration problem, the quadrature error can be split into  $m$  terms, as mentioned in Section 3.1

$$\|(\mathbf{I} - \mathbf{Q})u_{m,h}^p\| \lesssim \sum_{k=1}^m \|(I^{(k)} - Q^{(k)})u_{m,h}^p\|.$$

So there might be different error terms, depending on the problem and methods applied. Analogous to the curse of dimensionality described for the product quadrature, a small approximation error for  $\mathbb{E}[u^p(\mathbf{x})]$  requires a high truncation level, a fine discretization in space, and a large number of quadrature points, leading to large computational costs.

A possible approach to reduce the cost of computing the quantities of interest is the sparse grid method, which we will discuss in the following. It allows to break the curse of dimensionality, as it reduces the cost, while nearly preserving the accuracy.

In this and the next section, we will look at the sparse grid approach considering different parts of the problem. To be clear, we denote with  $d$  the number of numerical approximations taken into account for the sparse grid method, while the variable  $m$  is used for the stochastic dimensions. One objective will be to balance the spatial discretization and the quadrature in the stochastic parameters,

while keeping the truncation level fixed. Additionally, we will take into account a sparse grid also within the stochastic parameter set. In a later subsection, we will equilibrate the truncation error and quadrature error.

We follow an a-priori approach throughout and adapt methods based on knowledge about the regularity of the solution. First, we discuss the regular sparse grid method which treats each direction equally, and subsequently an extension to an anisotropic method. This section contains a short overview over works concerning sparse grid methods for stochastic PDEs, followed by the general concept of sparse grids. Then, we look at the sparse grid combination technique which allows to use the numerical methods discussed in Section 3. At the end, numerical examples are presented.

## 4.1 Sparse Grids

The approach of using sparse grids has been studied in different contexts. Basically, it can be applied whenever computational difficulties arise due to a high-dimensional tensor product structure of the problem with a related exponential growth of computational cost, and when the numerical methods applied allow for defining a sequence of methods with increasing approximation power.

The basic principle was first used by Smolyak for high-dimensional integration problems in the 1960s ([36]). When considering an  $m$ -fold product of one-dimensional nested quadrature rules, the approach reduces the number of quadrature points, which yields a substantial improvement in computational cost under the condition of additional smoothness of the integrand.

The sparse grid method was introduced to PDE problems by Zenger [37], whereby the expansion of a function in a hierarchical basis and the construction of associated subspaces based on tensor products of hierarchical increment spaces was developed. The sparse grid method has since been applied to various problems in diverse contexts, including problems from physics, chemistry, engineering and data mining (see [38] for more applications and references).

In the context of stochastic partial differential equations, Smolyak-type quadratures have been applied in stochastic collocation methods to reduce the computational cost of the quadrature in the parameter space for moderate dimensions ([19]), whereby anisotropic versions of the Smolyak quadrature have also been studied [21, 39].

In addition, a sparse grid approach between the spatial discretization and quadrature method for PDEs with a finite number of stochastic dimensions has been considered (see for example in [6, 23]). This combination is sometimes referred to as a multilevel method, as the sparse grid approach combining the spatial discretization and a Monte Carlo quadrature is related to the Multilevel Monte Carlo method. Similarly, other quadrature methods have been applied in place of a Monte Carlo method, including Quasi Monte Carlo and Smolyak-type quadratures (see, [12, 20, 40]). In [41], the perception of sparse grids as a multilevel

method has been extended to account for more than one spatial and one stochastic parameter.

Here, we consider stochastic PDEs where the diffusion coefficient is written in the Karhunen-Loeve expansion. In the literature, this is often referred to as uniformly elliptic case. The sparse grid has also been applied to problems with lognormal diffusion coefficients (e.g., [29, 42, 23, 43, 33]), However, we restrict the a-priori analysis to the uniformly elliptic case.

While a finite parameter space or fixed truncation level has often been studied in the above mentioned works, the interplay of the truncation error has also been investigated in some settings. The authors in [44] use a multilevel approach between Quasi Monte Carlo quadrature and an approximation to the PDE where the dimension is not fixed. Further research has been conducted looking at the computations of integrals over a countably infinite-dimensional domain (see, for example, [45, 46]).

In the following, the general concept of the sparse grid method is introduced. For that, we will follow [6] and use the overviews on sparse grids [38] and [47].

#### 4.1.1 General Concept

The sparse grid approach is based on a sequence of numerical methods and associated subspaces with increasing approximation power. An approximation space is constructed by combining increment spaces in an optimal way such that the approximation power is nearly preserved, but the dimension of the space does not grow exponentially. As the concept becomes clear when considering only two directions ( $d = 2$ ), we restrict ourselves to this case for the introduction of the sparse grid method. Later, the formulas for  $d > 2$  will be stated. We explain the sparse grid method in a general setting. Thus, the spaces and norms are not always specified. Nevertheless, we mainly focus on the case of a fixed stochastic dimension in this chapter, as a variable truncation level will be discussed separately later.

We start with sequences of finite-dimensional subspaces of some space  $\mathcal{H}_i$

$$V_0^{(i)} \subset V_1^{(i)} \subset \dots \subset \mathcal{H}_i \quad (17)$$

for  $i = 1, 2$ , such that the spaces have increasing approximation power and

$$\mathcal{H}_i = \overline{\bigcup_{l \in \mathbb{N}_0} V_l^{(i)}}. \quad (18)$$

In the context of the problem we are interested in, the spaces  $\{V_l^{(i)}\}$  refer to different sequences. If we are looking at a sparse grid involving the spatial discretization, the subspaces  $\{V_l^{(1)}\}$  are a sequence of approximation spaces for the spatial approximation, in our case finite element spaces, while  $\{V_l^{(2)}\}$  might correspond to a sequence of quadrature rules. For a sparse grid replacing a product

quadrature, both  $\{V_l^{(1)}\}$  and  $\{V_l^{(2)}\}$  correspond to sequences of one-dimensional quadrature rules which are stated sometimes in terms of the quadrature points involved. When considering the truncation level, the space corresponds to a function evaluation taking into account only finitely many variables.

The canonical way to obtain an approximation space in both directions would be combining the spaces in a tensor product

$$V_{l_1, l_2}^{\text{full}} = V_{l_1}^{(1)} \otimes V_{l_2}^{(2)},$$

to which we refer to as the full grid space for level  $L$  if  $l_1 = l_2 = L$ .

In order to obtain a sparse grid version of this tensor product, we define the increment spaces

$$W_{l+1}^{(i)} = V_{l+1}^{(i)} \setminus V_l^{(i)} \quad \text{with } W_0^{(i)} = V_0^{(i)}, \quad i = 1, 2.$$

From these, the full spaces can be reconstructed by

$$V_L^{(i)} = \bigotimes_{l=0}^L W_l^{(i)}, \quad i = 1, 2.$$

Hence, the full tensor product space  $V_L^{\text{full}}$  can be written as

$$V_L^{\text{full}} = \bigoplus_{l_1 \leq L, l_2 \leq L} W_{l_1}^{(1)} \otimes W_{l_2}^{(2)}.$$

Now, instead of the full sum, the regular sparse grid space only includes some increment spaces and is defined by

$$V_L^{\text{SG}} = \bigoplus_{l_1 + l_2 \leq L} W_{l_1}^{(1)} \otimes W_{l_2}^{(2)}. \quad (19)$$

The dimension of  $V_L^{\text{SG}}$  is considerably reduced compared to  $V_L^{\text{full}}$  if the dimensions of  $V_l^{(i)}$  form a geometric sequence, as then  $\dim V_L^{\text{SG}}$  is of order  $\mathcal{O}\left(\max\left\{\dim V_L^{(1)}, \dim V_L^{(2)}\right\}\right)$  up to logarithmic terms, whereas  $\dim V_L^{\text{full}} = \dim V_L^{(1)} \dim V_L^{(2)}$ .

For the computation of an approximation, using the sparse grid as an underlying approximation space implies the formula

$$\mathbf{P}_L^{\text{SG}} u = \sum_{l_1 + l_2 \leq L} \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} u, \quad (20)$$

where the difference operators  $\Delta_l^{(i)} : \mathcal{H}_i \rightarrow W_l^{(i)}$  are defined by

$$\Delta_l^{(i)} u = \begin{cases} \left(P_l^{(i)} - P_{l-1}^{(i)}\right) u & \text{if } l \geq 1 \\ P_l^{(i)} u & \text{if } l = 0. \end{cases}$$

The operators  $P_l^{(i)}$  correspond to applying the numerical method of level  $l$ . For example, it can stand for a quadrature method with  $N_l$  nodes. In the case where a spatial approximation to  $u^p$ ,  $p = 1, 2$ , is considered, the operators  $P_l^{(2)}$  are given by

$$P_l^{(2)}u = \mathcal{F}(u_{h_l}) := u_{h_l}^p \quad p = 1, 2, \quad (21)$$

where  $u_{h_l}$  is the projection of  $u$  onto the finite element space of level  $l$  computed by the finite element method.

In contrast to (20), the full grid solution is

$$\mathbf{P}_L^{\text{full}}u = \sum_{l_1 \leq L, l_2 \leq L} \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)}u.$$

Next, we want to motivate the choice of summing over the set  $\{(l_1, l_2) : l_1 + l_2 \leq L\}$  in the sparse grid formulation. We refer to [38] for a detailed derivation. Using (18), we can write the exact solution  $\mathbf{P}_\infty u$  in the expansion

$$\mathbf{P}_\infty u = \sum_{l_1, l_2=0}^{\infty} \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)}u.$$

The idea of the sparse grid is to find an approximation by only summing over a finite set of indices  $\mathcal{I} \subset \mathbb{N}^2$

$$\mathbf{P}_{\mathcal{I}}u = \sum_{(l_1, l_2) \in \mathcal{I}} \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)}u.$$

Hence, the error of this approximation can be bounded by

$$\|\mathbf{P}_\infty u - \mathbf{P}_{\mathcal{I}}u\| \leq \sum_{(l_1, l_2) \notin \mathcal{I}} \|\Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)}u\|.$$

The aim now is to find, for a prescribed computational cost, the optimal index set  $\mathcal{I}$  such that the error of the approximation is minimal. This can be formulated as an optimization problem, referred to as a knapsack problem in combinatorial optimization (see [38] for details).

A solution to the knapsack problem can be found, by considering the benefit-cost ratio

$$\text{bcr}(l_1, l_2) = \frac{\|\Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)}u\|}{\text{cost}(\Delta_{l_1}^{(1)}) \text{cost}(\Delta_{l_2}^{(2)})},$$

where  $\text{cost}(\Delta_{l_i}^{(i)})$  denotes the cost of the evaluation  $\Delta_{l_i}^{(i)}u$ . The optimal set  $\mathcal{I}$  then only includes the indices  $\mathbf{l} = (l_1, l_2)$  with the highest benefit-cost ratios. Indices with a benefit-cost ratio below a threshold parameter are not used in the sparse grid.

In the case where  $\|\Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} u\| \lesssim 2^{-C(l_1+l_2)}$  and also  $\text{cost}(\Delta_{l_1}^{(1)}) \text{cost}(\Delta_{l_2}^{(2)}) \approx 2^{C'(l_1+l_2)}$  for constants  $C, C' > 0$ , the index set takes the form

$$\mathcal{I} = \{(l_1, l_2) \in \mathbb{N}^2 \mid l_1 + l_2 \leq L\}. \quad (22)$$

However, if the benefit-cost ratio has a different structure, the optimal set looks different. For example, we will look at a version of the sparse grid where the benefit-cost ratio is of the form  $\text{bcr}(l_1, l_2) \leq 2^{-(s_1+n_1)l_1 - (s_2+n_2)l_2}$  with possibly  $s_1 + n_1 \neq s_2 + n_2$  in Section 5. Often in Multilevel methods, the numerical methods are combined such that the accuracy is balanced even though the cost required is not equal. In this case, the regular sparse grid method with set  $\{(l_1, l_2) \in \mathbb{N}^2 \mid l_1 + l_2 \leq L\}$ , which balances the approximation powers, also reduces the cost, but not in an optimal way.

We will come back to the concept of benefit-cost ratio, when constructing an adaptive algorithm that determines a good index set  $\mathcal{I}$  in Section 6. For completeness, we state the generalization of the sparse grid method for more than two dimensions. For  $d > 2$ , the regular sparse grid space is defined as

$$V_L^{\text{SG}} = \bigoplus_{|l_1| \leq L} W_{l_1}^{(1)} \otimes W_{l_2}^{(2)} \dots \otimes W_{l_d}^{(d)}$$

and the approximation is obtained by the formula

$$P_L^{\text{SG}} u = \sum_{|l_1| \leq L} \Delta_l u,$$

where

$$\Delta_l u = \left( \Delta_{l_1}^{(1)} \otimes \dots \otimes \Delta_{l_d}^{(d)} \right) u.$$

Regarding the computational cost of the sparse grid, let us assume that  $\dim W_{l_i}^{(i)} = 2^{l_i}$  for  $i = 1, \dots, d$ . Then the dimension of the sparse grid space is of order

$$\mathcal{O}(N_L \log_2(N_L)^{d-1}),$$

where  $N_L = 2^L$  (see [48]), which is substantially smaller than  $N_L^d$ , the dimension of the full grid space. As the exponential dependence on  $d$  is only in the logarithmic term, the dimension of the sparse grid space is almost reduced to a linear term. Despite the considerable reduction in the dimension of the approximation space, the approximation power is not affected substantially, if additional regularity is assumed. The precise conditions will be discussed in the next section.

#### 4.1.2 Mixed Regularity

For the sparse grid approach to not deteriorate the accuracy, the contribution of the indices that have been left out should be small. Therefore, we require a decrease of the contribution with increasing levels, e.g.,

$$\|\Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} u\| \lesssim 2^{-C(l_1+l_2)}. \quad (23)$$

This condition is satisfied when the function that is approximated fulfills some mixed regularity conditions. In this subsection, we discuss the regularity assumption which ensures that the accuracy is preserved for the sparse grid in the case of a Smolyak quadrature and for a sparse grid between spatial and stochastic approximation, with a fixed stochastic dimension.

With regard to a sparse grid in the quadrature, the integrand needs to have mixed derivatives up to order  $r$ . More precisely, it needs to lie in the space

$$u \in \left\{ v : \Gamma \rightarrow \mathbb{R}, \left\| \frac{\partial^{|\alpha|} v}{\partial y_1^{\alpha_1} \dots \partial y_m^{\alpha_m}} \right\|_{L^\infty(\Gamma)} < \infty \text{ for } \alpha_i \leq r \right\},$$

where the order  $r$  depends on the choice of the one-dimensional quadrature rules.

Under this assumption, the convergence rate with respect to the computational cost is given by

$$\mathcal{O}(N_L^{-r} \log(N_L)^{(d-1)(r+1)}) \quad (24)$$

(see [49]), which is the rate of the one-dimensional quadrature, reduced by only a logarithmic term.

Including the spatial direction in the sparse grid, the mixed regularity needs to hold additionally between the  $x$  and  $y$  variables. The condition (23) between the spatial and stochastic approximation takes the form

$$\| (\mathbf{Q}_{l_1} - \mathbf{Q}_{l_1-1}) \left( \mathcal{F}(u_{h_{l_2}}) - \mathcal{F}(u_{h_{l_2-1}}) \right) \|_{\mathcal{X}} \lesssim 2^{-(l_1+l_2)} \|f\|_{L^2(D)}^p, \quad (25)$$

where the function  $\mathcal{F}$  is as in (21). Depending on whether the first or second moment is computed,  $\mathcal{X} = H^1(D)$  for  $p = 1$ , or  $\mathcal{X} = W^{1,1}(D)$  for  $p = 2$ . This condition needs to be checked separately for the different quadrature rules, as the quadratures require different regularity assumptions on the integrand.

If  $\mathbf{Q}$  is a Monte Carlo quadrature, the condition

$$\| (\mathbf{Q}_{l_1} - \mathbf{Q}_{l_1-1}) \left( \mathcal{F}(u_{h_{l_2}}) - \mathcal{F}(u_{h_{l_2-1}}) \right) \|_{L_\rho^2(\mathcal{X})} \lesssim 2^{-(l_1+l_2)} \|f\|_{L^2(D)}^p$$

is straightforward and is satisfied for an integrand that is square integrable with respect to  $\mathbf{y}$  ([12]). The norm in (25), however, needs to be replaced by  $L_\rho^2(\Gamma, \mathcal{X})$  as the quadrature does not provide a deterministic error analysis.

For other quadrature rules, the analysis of the error is based on estimates for the derivatives of the integrand. Therefore, the derivatives of  $(\mathcal{F}(u_{h_l}) - \mathcal{F}(u_{h_{l-1}}))$  need to be bounded. The following results have been proven in [6] for the derivatives of the difference between approximate and exact solution to the PDE.

**Lemma 5.** For the PDE (3) with diffusion coefficient given by the truncated Karhunen-Loeve expansion, it holds for all  $\boldsymbol{\alpha} \in \mathbb{N}^m$  and  $\mathbf{y} \in \Gamma$  that

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}(u - u_{h_l})(\mathbf{y})\|_{H^1(D)} \lesssim 2^{-l} |\boldsymbol{\alpha}|! c^{|\boldsymbol{\alpha}|} \boldsymbol{\gamma}^{\boldsymbol{\alpha}} \|f\|_{L^2(D)}$$

and

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}(u^2 - u_{h_l}^2)(\mathbf{y})\|_{W^{1,1}(D)} \lesssim 2^{-l} |\boldsymbol{\alpha}|! c^{|\boldsymbol{\alpha}|} \boldsymbol{\gamma}^{\boldsymbol{\alpha}} \|f\|_{L^2(D)}^2,$$

where  $\boldsymbol{\gamma}$  is related to the coefficient in the Karhunen-Loeve expansion and the hidden constants are independent of  $\mathbf{y}$ .

With these estimates, the required mixed regularity (25) can be derived for the quadrature rules different from Monte Carlo (see [6, 15]), whereby also the mixed regularity between the stochastic variables is provided.

Based on (25), the following convergence estimate with respect to the sparse grid level was proven in [6] for an equivalent representation of the sparse grid which will be discussed in the next section.

**Theorem 4.** The computation of  $\mathbb{E}[\mathcal{F}(u)]$  with a sparse grid approach under condition (25) satisfies

$$\|\mathbb{E}[\mathcal{F}(u)] - \sum_{l_1+l_2 \leq L} (\mathbf{Q}_{l_1} - \mathbf{Q}_{l_1-1})(\mathcal{F}(u_{h_{l_2}}) - \mathcal{F}(u_{h_{l_2-1}}))\|_{\mathcal{X}} \lesssim 2^{-L} L \|f\|_{L^2(D)}^p.$$

## 4.2 Combination Technique

In this section, we focus on a different representation of the sparse grid, which is also known as the combination technique and was introduced in [50]. Instead of considering the increment spaces  $W_l$ , the combination technique relies on a reformulation of the sparse grid formula (20) and combines solutions in certain full, but small, tensor product spaces  $V_l$ . Since the combination technique is equivalent to the sparse grid formula, the same error bounds hold.

Again, we will consider for simplicity the case  $d = 2$  first. The underlying idea of the combination technique is to use the telescope sum property to recover a representation based on full grid solutions. Rearranging the terms in (20) yields

$$\begin{aligned} \mathbf{P}_L^{\text{SG}} u &= \sum_{l_1+l_2 \leq L} (\mathbf{P}_{l_1, l_2} - \mathbf{P}_{l_1, l_2-1} - \mathbf{P}_{l_1-1, l_2} + \mathbf{P}_{l_1-1, l_2-1}) u \\ &= \sum_{l_1=0}^L \sum_{l_2=0}^{L-l_1} \left( P_{l_1}^{(1)} \left( P_{l_2}^{(2)} - P_{l_2-1}^{(2)} \right) u - P_{l_1-1}^{(1)} \left( P_{l_2}^{(2)} - P_{l_2-1}^{(2)} \right) u \right) \\ &= \sum_{l_1=0}^L P_{l_1}^{(1)} P_{L-l_1}^{(2)} u - \sum_{l_1=0}^L P_{l_1-1}^{(1)} P_{L-l_1}^{(2)} u \\ &= \sum_{l_1+l_2=L} \mathbf{P}_{l_1, l_2} u - \sum_{l_1+l_2=L-1} \mathbf{P}_{l_1, l_2} u \\ &=: \mathbf{P}_L^{\text{ct}} u. \end{aligned} \tag{26}$$

In higher dimensions the combination technique representation is given by

$$\mathbf{P}_L^{\text{ct}} u = \sum_{i=0}^{d-1} (-1)^i \binom{d-1}{i} \sum_{|\mathbf{l}|_1=L-i} \mathbf{P}_{\mathbf{l}} u. \quad (27)$$

It is also possible to provide a formula in terms of full grid solutions for a generic index set  $\mathcal{I} \subset \mathbb{N}_0^d$  which satisfies some admissibility condition. As we will only use it in Section 6, we do not specify the condition here and state only the formula for completeness:

$$\mathbf{P}_{\mathcal{I}}^{\text{ct}} u = \sum_{\mathbf{l} \in \mathcal{I}} \left( \sum_{\mathbf{z}=\mathbf{0}}^{\mathbf{1}} (-1)^{|\mathbf{z}|_1} \chi^{\mathcal{I}}(\mathbf{l} + \mathbf{z}) \right) \mathbf{P}_{\mathbf{l}} u,$$

where  $\mathbf{1} = (1, \dots, 1)$  and  $\chi^{\mathcal{I}}(\mathbf{l}) = \begin{cases} 1 & \text{if } \mathbf{l} \in \mathcal{I}, \\ 0 & \text{otherwise.} \end{cases}$

The combination technique has some advantages compared to the original sparse grid formulation. First, the formula (27) includes only full grid solutions  $\mathbf{P}_{\mathbf{l}} u$  and it does not require the computation of increments  $\Delta_{\mathbf{l}} u$ . This has the major advantage that no solvers on the increment spaces  $W_{\mathbf{l}}$  need to be available, which might be difficult to construct. Instead, standard numerical methods can be applied. In addition, the use of full grid spaces allows dropping the assumption of nested approximation spaces in (17). Hence, we do not necessarily need to consider quadrature methods and finite element spaces that are nested.

Furthermore, the structure of (27) allows performing parallel computations. As the computation of the full grid solutions  $\mathbf{P}_{\mathbf{l}} u$  are independent from each other, they can be solved in parallel. In this way, the computational time can be further reduced. Because of these advantages, we will use the combination technique representation of the sparse grid in our implementations.

Regarding the computational complexity of the combination technique, it can be shown that it is of the same order as for the sparse grid. The combination technique requires solving  $\mathcal{O}(L^{d-1})$  problems in small full grid spaces and the computational cost is dominated by computing  $\mathbf{P}_{\mathbf{l}} u$  for  $|\mathbf{l}|_1 = L$ , for which there are  $\binom{L+d-1}{d-1}$  combinations to form this sum. Hence, if the cost for computing the solution in one direction is  $N_l^{(i)} = 2^l$ , then

$$\text{cost}(\mathbf{P}_L^{\text{ct}}) = \mathcal{O}(N_L \log_2 N_L^{d-1}).$$

In addition to the combination technique formula (27), other representations of the sparse grid have been used which consider a mixture of full and increment spaces (see [6]). For example, the intermediate step (26) can be considered such that

$$\mathbf{P}_L^{\text{SG}} u = \sum_{l=0}^L \Delta_l^{(1)} P_{L-l}^{(2)} u. \quad (28)$$

By summing over  $l_1$  first, the roles  $l_1$  and  $l_2$  are reversed such that

$$\mathbf{P}_L^{\text{SG}} u = \sum_{l=0}^L P_{L-l}^{(1)} \Delta_l^{(2)} u. \quad (29)$$

The corresponding spaces are given by

$$V_L^{\text{SG}} = \bigoplus_{l_1=0}^L W_{l_1}^{(1)} \otimes \left( \bigoplus_{l_2=0}^{L-l_1} W_{l_2}^{(2)} \right) = \bigoplus_{l=0}^L W_l^{(1)} \otimes V_{L-l}^{(2)}$$

and by

$$V_L^{\text{SG}} = \bigoplus_{l_2=0}^L \left( \bigoplus_{l_1=0}^{L-l_2} W_{l_1}^{(1)} \right) \otimes W_{l_2}^{(2)} = \bigoplus_{l=0}^L V_{L-l}^{(1)} \otimes V_l^{(2)}.$$

It is convenient to use this representation when it is easy to construct a nested sequence of finite-dimensional approximation spaces for one direction but not the other.

The formulas (28) and (29) also demonstrate the connection to Multilevel methods used in [12, 51], which originate from the numerical integration of stochastic differential equations [52] and apply different quadrature levels to a sequence of differences of spatial discretizations. For the coarsest spatial approximation, the most accurate quadrature is used, while for the finest spatial approximation, the least accurate quadrature is applied. Basically, the multilevel method employs the representation (29) with the sequence  $\{P_l^{(1)}\}_l$  corresponding to a sequence of quadrature rules, and  $\{P_l^{(2)} u\}_l$  to the approximation to the PDE.

### 4.3 Numerical examples

#### Inclusion Problem

We first illustrate the advantages of the sparse grid approach compared to the product quadrature. For that, we look at a problem where the sparse grid is applied to the integration only. We adopt the inclusion problem studied in [53], and compute the mean of the solution to the PDE, where the diffusion coefficient depends on four stochastic parameters. The inclusion problem is derived from a thermal conduction problem where the conductivity is random in disjoint circular subdomains. More precisely, the diffusion coefficient is given by

$$a(\mathbf{x}, \mathbf{y}) = a_0 + \sum_{k=1}^4 \chi_{D_k}(\mathbf{x}) y_k,$$

where  $\chi_{D_k} = 1$  on the set  $D_k \subset D$  and zero elsewhere. The subdomains  $D_k$  are circles centered at  $\mathbf{x}_k$  with radius  $\frac{1}{8}$ , i.e.,  $D_k = B(\mathbf{x}_k, \frac{1}{8})$ , with

$$\{\mathbf{x}_k\}_{k=1}^4 = \left\{ \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}, \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 0.8 \\ 0.8 \end{pmatrix} \right\}.$$

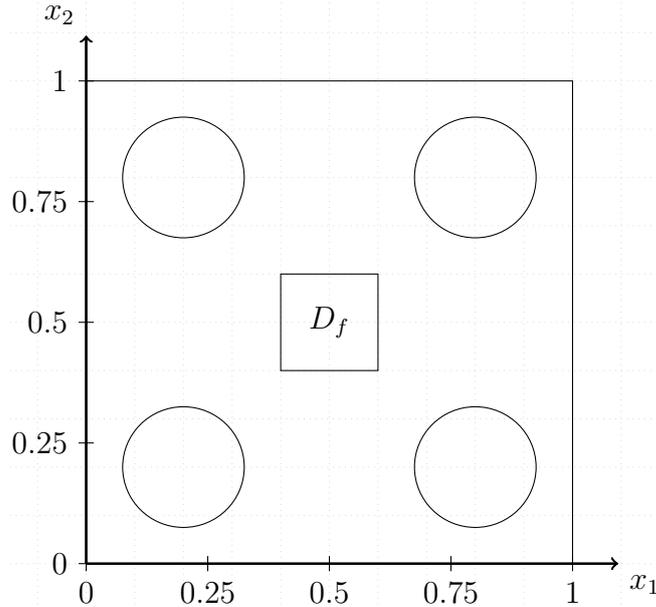


Figure 1: Illustration of the sets used in the inclusion problem.

For the right hand side of the PDE, the function  $f(\mathbf{x}) = 100\chi_{D_f}(\mathbf{x})$  is applied with  $D_f = [0.4, 0.6]^2 \subset D$ . An illustration of this setup is shown in Figure 1.

Furthermore, the parameters  $y_k$  are uniformly distributed in  $[-0.99, -0.2]$  and we set  $a_0 = 1$  which ensures the coercivity of the coefficient. Due to the structure of  $a(\mathbf{x}, \mathbf{y})$ , the influence of the stochastic variables is equally strong. We apply the combination technique with trapezoidal quadrature rules for each of the four stochastic dimensions while keeping the spatial discretization fixed with a mesh of  $N_x$  grid points in each spatial dimension, where  $N_x = 64$ .

The convergence behavior is depicted in Figure 2. Choosing  $N_l^{(i)} = 2^l$  quadrature points for the trapezoidal rules, we expect a rate of the error of  $\mathcal{O}(2^{-2L}L^{d-1})$ . This rate can be observed, although the logarithmic term is almost not visible. Furthermore, we compare the convergence of the combination technique with respect to the computational cost with the full grid product quadrature. The plots include the theoretical convergence rates for comparison. We observe that the convergence rate of the trapezoidal quadrature with  $\mathcal{O}(N^{-2})$  is almost preserved with the combination technique and is only deteriorated by a logarithmic term. Asymptotically, this logarithmic term will have less impact such that the rate approaches  $\mathcal{O}(N^{-2})$ , whereas the rate for the full grid is  $\mathcal{O}(N^{-1/2})$  due to the dimensionality of the integration. This illustrates the curse of dimensionality and the advantages of the sparse grid in this context.

So far, the spatial discretization has been fixed. From the convergence, we see however, that we can obtain relatively small errors with the combination technique in the stochastic variables. A balance of the finite element error would require a high discretization in space. We therefore consider from now on numerical examples for the combination technique with respect to both the finite element method and quadrature rule.

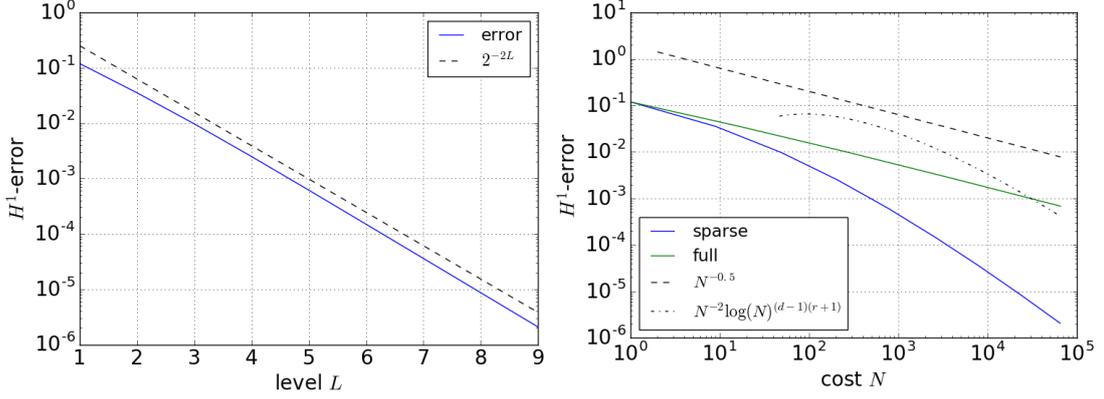


Figure 2: Convergence of the combination technique for the computation of  $\mathbb{E}[u(\mathbf{x})]$  applied to the stochastic parameters measured with respect to a full grid reference solution. Left: convergence with respect to the sparse grid level, Right: convergence with respect to the computational cost compared to a full grid product quadrature.

### Spatial and Stochastic Discretization

First, we consider a constructed example with one spatial and one stochastic variable. We look at the model problem for  $y \in \mathcal{U}(-1, 1)$

$$\begin{aligned} -\operatorname{div}(a(x, y)\nabla u(x, y)) &= f(x) & \text{in } D = (-1, 1) \\ u(x, y) &= 0 & \text{on } \partial D \end{aligned}$$

with

$$a(x, y) = e^{-y} \cos\left(\frac{\pi x}{2}\right) \quad \text{and} \quad f(x) = \frac{\pi^2}{4} \cos(\pi x) \quad (30)$$

such that the solution to the PDE is known and has the form

$$u(x, y) = e^y \cos\left(\frac{\pi x}{2}\right).$$

Hence, the function is smooth in both variables  $x$  and  $y$  and the regularity assumptions are satisfied, which are required for applying the sparse grid between the space and stochastic variable for all three quadrature rules.

For example the trapezoidal rule for the computation of  $\mathbb{E}[u(\mathbf{x})]$  can be used. Choosing for the finite element method of level  $l$  a mesh with  $N_l^x = 2^l$  points, we have by Lemma 5 for all  $y \in [-1, 1]$

$$\|\partial_y^2(u - u_{h_l})(y)\|_{H^1(D)} \lesssim 2^{-l} \|f\|_{L^2(D)}^2.$$

If we also apply the quadrature with  $N_l^y = 2^l$  quadrature points, the convergence of the trapezoidal rule (6) yields

$$\|(\mathcal{Q}_{l_1} - \mathcal{Q}_{l_1-1})(u_{h_{l_2}} - u_{h_{l_2-1}})\|_{H^1(D)} \lesssim 2^{-(2l_1+l_2)} \|f\|_{L^2(D)},$$

whereby the cost is of order  $\mathcal{O}(2^{l_1+l_2})$ .

In view of the convergence rate of the finite element method with respect to the  $L^2$ -norm, we can expect to achieve the same rate of convergence in the finite element method as for the quadrature methods such that

$$\|(\mathcal{Q}_{l_1} - \mathcal{Q}_{l_1-1})(u_{h_{l_2}} - u_{h_{l_2-1}})\|_{L^2(D)} \lesssim 2^{-(2l_1+2l_2)} \|f\|_{L^2(D)}.$$

We plot the benefit-cost ratio in Figure 3. As expected, the benefit-cost ratio becomes small when both discretization levels  $l_1$  and  $l_2$  become large. A triangular shape of the set which consists of indices with a benefit-cost ratio greater than a threshold, is clearly visible. Furthermore, we observe that the the optimal index set with respect to the  $L^2(D)$ -norm includes the discretization methods up to the same level. Hence, the regular sparse grid with the index set (22) is the optimal choice. In comparison to that, the triangle for the error measured in the  $H^1(D)$ -norm includes higher discretization levels for the spatial variable than for the stochastic, because the convergence of the spatial method is slower. Applying a regular sparse grid yields, therefore, still an improvement. However, it is not the optimal choice. A better choice for the index set will be discussed in the next section.

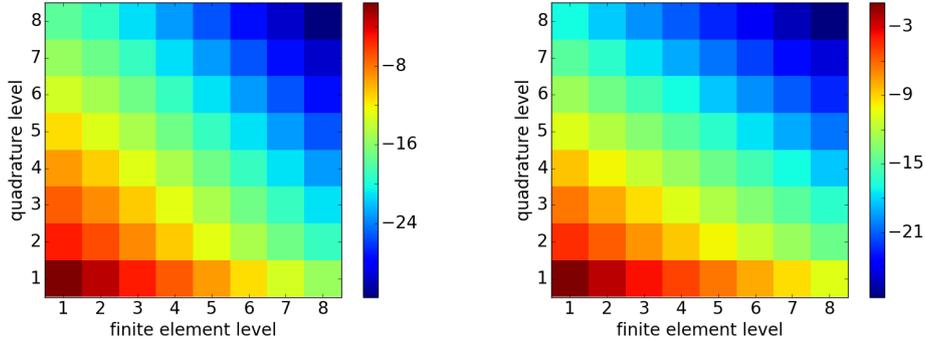


Figure 3: Benefit-cost ratio for the finite element discretization and trapezoidal rule. The plots are with a logarithmic scale and the benefit is measured in the  $L^2(D)$ -norm (left) and  $H^1(D)$ -norm (right).

We apply the combination technique to this example problem using the Monte Carlo, trapezoidal and Gauss-Legendre rules. Since the convergence rates of the stochastic method in terms of quadrature points do not coincide with the rate of the finite element method, the regular sparse grid is not optimal. Nevertheless, we can equilibrate the errors and obtain the convergence rate of Theorem 4.

Moreover, we use a finite element method with  $N_l^x = 2^l$  such that the error in the  $H^1(D)$ -norm is of order  $\mathcal{O}(2^{-l})$  and for the quadrature the number of points  $N_l$  such that the error is of the same order. In addition, we include the convergence plots of the combination technique for the trapezoidal rule balanced to the PDE

solver in the  $L^2(D)$ -norm. In this case the spatial and stochastic approximation have the same convergence rate and the benefit-cost ratio is of the form which was assumed for the derivation of the regular sparse grid.

The convergence plots are shown in Figure 4. In terms of the sparse grid level, the error behaves according to the error of the PDE discretization up to a logarithmic term. This is consistent with the error bound in Theorem 4, but we observe a factor  $\sqrt{L}$  instead of  $L$ . Also, in this example the curse of dimensionality is broken, as the convergence with respect to the cost is the lower rate of the two methods, which is only affected by a logarithmic factor.

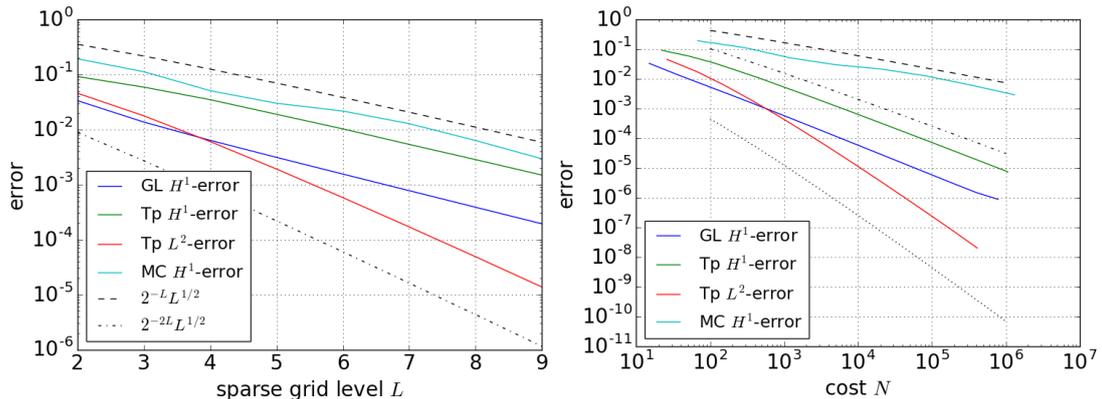


Figure 4: Convergence results for the regular sparse grid for the Gauss-Legendre (GL), trapezoidal (Tp) and Monte Carlo (MC) quadrature method. Left: convergence with respect to sparse grid level. Right: convergence with respect to the computational cost. As a reference the rates  $\mathcal{O}(N^{-1/2} \log(N)^{1/2})$ ,  $\mathcal{O}(N^{-1} \log(N)^{1/2})$  and  $\mathcal{O}(N^{-2} \log(N)^{5/2})$  are shown.

### PDE with Random Field as Diffusion Coefficient

As a third example, we look at a situation where the diffusion coefficient is given by a random field and the Karhunen-Loeve expansion needs to be truncated. We consider the PDE

$$\begin{aligned} -\operatorname{div}(a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) &= 1 && \text{in } D = (0, 1) \\ u(\mathbf{x}, \omega) &= 0 && \text{on } \partial D \end{aligned}$$

together with the random field with  $\mathbb{E}[a(\mathbf{x})] = 10$  and Gaussian covariance function

$$\operatorname{cov}_a(\mathbf{x}_1, \mathbf{x}_2) = 2 \exp\left(-\frac{1}{2} \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{0.1}\right),$$

for which we compute the Karhunen-Loeve expansion. As the covariance function is smooth with a relatively long correlation length, only few terms of the Karhunen-Loeve expansion need to be considered. We fix the level of truncation to eight stochastic variables and apply the combination technique in space and

stochastic, whereby we use one-dimensional Gauss-Legendre quadrature rules. By Lemma 5 and Lemma 2, the function  $u$  can be extended into some region of the complex plane, such that exponential convergence of the Gauss-Legendre quadrature holds. Furthermore, these statements provide for an anisotropy of the stochastic variables expressed in the sequence  $\gamma_k$ . We use this to choose the quadrature points such that the error is of order  $\mathcal{O}(2^{-l})$ .

For this example, we also consider, in addition to the computation of the expectation  $\mathbb{E}[u]$ , the second moment  $\mathbb{E}[u^2]$ . The  $W^{1,1}(D)$ -norm is not accessible in the finite element software. We therefore look at the error in the  $H^1(D)$ -norm for the second moment, as  $H^1(D) \subset W^{1,1}(D)$  is continuous embedded and therefore  $\|v\|_{W^{1,1}(D)} \leq C(D)\|v\|_{H^1(D)}$ .

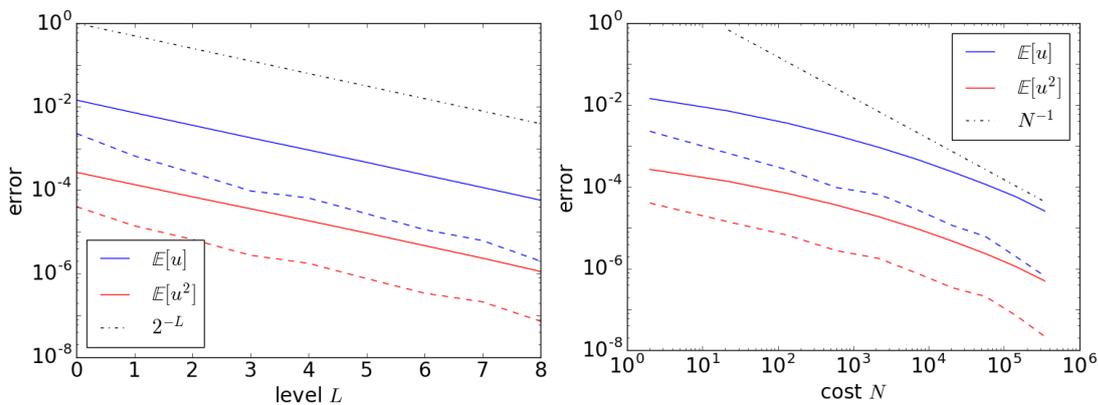


Figure 5: Convergence behavior of the computation of first and second moment. Errors are measured in  $H^1(D)$  (solid) and  $L^2(D)$  (dashed).

The convergence of the combination technique with eight stochastic and one spatial direction is shown in Figure 5. According to the choice of approximation power of the methods, we expect a convergence rate of essentially  $\mathcal{O}(2^{-L})$  with respect to the level  $L$  and that the rate of  $\mathcal{O}(N^{-1})$  with respect to the computational cost  $N$  is deteriorated at most by a logarithmic factor. Indeed, we observe the expected behavior, whereby the powers on the logarithmic factors are smaller than suggested. This is due to the fact that for small levels the number of quadrature points used is larger than required for a rate  $\mathcal{O}(2^{-l})$ , as the number of points needs to be an integer. Furthermore, the number of quadrature points does not grow geometrically with the level, which leads to a slower growth in the computational cost than assumed for (24).

In addition to the  $H^1(D)$ -error, we plot the  $L^2(D)$ -error of the computation, as it shows how the convergence rate of the sparse grid is dominated by the slowest method. From the finite element method, we could expect that the error decreases with order 2. However, since we do not achieve this rate with the quadrature methods, it is reduced to the slower convergence.

## 5 Anisotropic Sparse Grids

As seen in the numerical examples, the regular sparse grid does not always reflect the behavior of the benefit-cost ratio. In the case where we apply numerical methods with different convergence rates for the directions of the sparse grid, it is favorable to use an anisotropic version of the sparse grid approach instead. For example, if the method used in one direction converges faster than the method in the second direction, the convergence is dominated by the slower method and less computational work should be considered for the faster method. Therefore, we will discuss in this section an approach that does not treat each direction equally and that allows control of this anisotropy with a parameter. The parameter can be chosen such that the methods are balanced based on an equal amount of work, equal approximation power or equilibrated benefit-cost ratio.

The anisotropic sparse grid was first described in [5]. There, an anisotropic sparse grid was introduced for the approximation of a function on a product domain  $\Omega_1 \times \Omega_2$ , where  $\Omega_1 \subset \mathbb{R}^{n_1}$  and  $\Omega_2 \subset \mathbb{R}^{n_2}$ . For an approximation of this function in  $L^2(\Omega_1 \times \Omega_2)$ , the method combines finite-dimensional subspaces with different approximation power and dimensions in an optimal way. This method was also transferred to the problem of estimation of an econometric model in [54].

The anisotropic sparse grid is basically an extension of the regular sparse grid as described in Section 4 in the sense that it adjusts the index set  $\mathcal{I}$  used for the sparse grid based on the cost and accuracy properties of the numerical methods.

While the regular sparse grid uses the product of increment spaces  $W_{l_1} \otimes W_{l_2}$  with the indices satisfying  $l_1 + l_2 \leq L$ , which, viewed as a set in the  $l_1, l_2$ -plane, form a triangular shaped domain with equally sized sides, the set of indices in the anisotropic formulation form a scalene triangle. This allows including more subspace in one direction than in the other. The particular shape of the scalene triangle is determined by a parameter  $\sigma > 0$  in such a way that the ratio of the legs of the triangle is given by  $\sigma^2$ . This is reflected in the condition  $l_1\sigma + \frac{l_2}{\sigma} \leq L$  for the indices. Accordingly, the anisotropic sparse grid space in terms of incremental spaces is defined by

$$V_L^\sigma = \sum_{\sigma l_1 + l_2/\sigma \leq L} W_{l_1} \otimes W_{l_2}.$$

The difference between the regular and anisotropic sparse grid is illustrated in Figure 6.

In the following, the details of the anisotropic sparse grid together with the convergence properties will be presented for combining two different numerical methods. We adapt the ideas of [5] and first consider the interplay between the spatial and stochastic method. In a subsequent part, the anisotropic sparse grid approach is used to balance the truncation of an infinite-dimensional parameter space and its stochastic approximation.

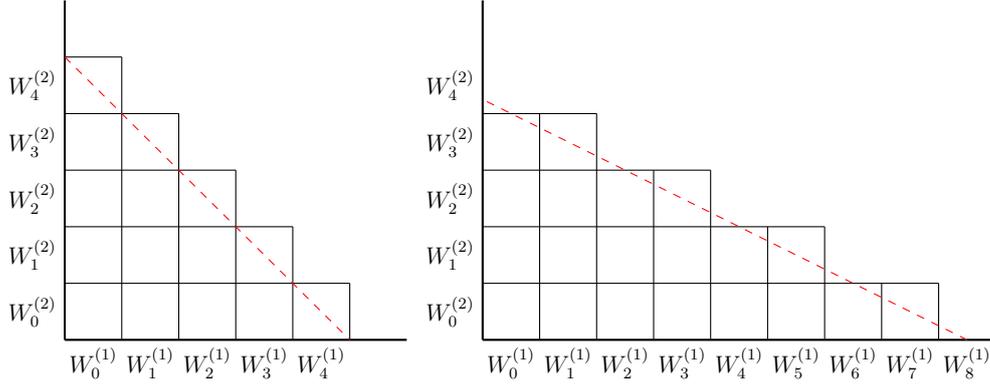


Figure 6: Subspaces included in the sparse grid formulation. Left: regular sparse grid, right: anisotropic sparse grid.

## 5.1 Anisotropic Sparse Grid Method for Quadrature and Spatial Discretization

First, we discuss an anisotropic version of the sparse grid for the combination of spatial and stochastic discretization. As before, we consider sequences of numerical methods, which correspond to a sequence of quadrature rules  $\{Q_l\}_{l \geq 0}$  and the function  $\mathcal{F}$  applied to a sequence of projections onto finite element spaces, i.e.,  $\{\mathcal{F}(u_{h_l})\}_{l \geq 0}$ , where  $\mathcal{F}(u) = u$  or  $\mathcal{F}(u) = u^2$ .

We require again that the approximation increases such that

$$\lim_{l \rightarrow \infty} Q_l v = \int v(z) dz \quad \text{and} \quad \lim_{l \rightarrow \infty} \mathcal{F}(u_{h_l}) = \mathcal{F}(u)$$

and define as before the difference operator

$$\Delta_l^{(1)} v = \begin{cases} Q_l v - Q_{l-1} v & \text{if } l \geq 1, \\ Q_0 v & \text{if } l = 0, \end{cases} \quad (31)$$

and

$$\Delta_l^{(2)} u = \begin{cases} \mathcal{F}(u_{h_l}) - \mathcal{F}(u_{h_{l-1}}) & \text{if } l \geq 1, \\ \mathcal{F}(u_{h_0}) & \text{if } l = 0. \end{cases} \quad (32)$$

The difference in approximation power and cost of the spatial and stochastic discretization is described with two parameters. We assume that the number of quadrature points for  $Q_l$  is given by  $\text{cost}(Q_l) \approx 2^{n_1 l}$  and that the degrees of freedom of the finite element spaces are chosen such that  $\text{cost}(u_{h_l}) \approx 2^{n_2 l}$ . The accuracy of the methods are described by  $s_1$  and  $s_2$ , as we assume that the error is of order  $\mathcal{O}(2^{-s_i l})$  for  $i = 1, 2$ .

In [5],  $n_i$  refers to the dimension of the underlying spaces and  $s_i$  to the degree of polynomial exactness of the subspaces. Nevertheless, it is more convenient to think of  $n_i$  as a parameter that describes the cost needed for an error of

order  $\mathcal{O}(2^{-s_i l})$ . Hence, it is not necessarily a positive integer. The values of  $n_i$  and  $s_i$  depend on the applied numerical method. However, the choice of  $n_i$  implies the value of  $s_i$ , as the ratio  $s_i/n_i$  is fixed.

As for the regular sparse grid method, we need to require mixed regularity in order for the sparse grid to work. We assume

$$\left\| \left( \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} \right) u \right\|_{\mathcal{X}} \lesssim 2^{-(s_1 l_1 + s_2 l_2)} \|f\|_{L^2(D)}^p, \quad (33)$$

which incorporates the approximation powers  $s_1$  and  $s_2$ .

Provided these assumption on the numerical methods and regularity of  $u$ , the anisotropic sparse grid solution is defined as

$$\mathbf{Q}_L^\sigma u = \sum_{\sigma l_1 + l_2 / \sigma \leq L} \left( \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} \right) u, \quad (34)$$

where the choice of  $\sigma$  depends on what we want to equilibrate. If  $\sigma$  is chosen according to  $\sigma = \sqrt{n_1/n_2}$ , the computational cost of both methods on the indices  $\sigma l_1 + l_2 / \sigma = L$  is balanced. An equilibrate error rate is obtained with  $\sigma = \sqrt{s_1/s_2}$ , whereas an equilibrated benefit-cost ratio is accomplished with

$$\sigma = \sqrt{\frac{n_1 + s_1}{n_2 + s_2}}.$$

Before discussing the computational cost and error rate for (34), we consider a reformulation of (34) similar to the combination technique (see Section 4.2). This has the advantage that the increments  $\Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} u$  do not need to be computed, instead we can use full combinations of the quadrature and spatial discretization  $Q_{l_1} \mathcal{F}(u_{h_{l_2}})$ .

For that, we note that the condition  $l_1 \sigma + l_2 / \sigma \leq L$  implies

$$\begin{aligned} 0 \leq l_1 &\leq \frac{1}{\sigma} L - \frac{1}{\sigma^2} l_2, \\ 0 \leq l_2 &\leq \sigma L - \sigma^2 l_1. \end{aligned}$$

Using the telescoping sum, we obtain the following formulations. For  $\sigma \geq 1$  it is simpler to use

$$\mathbf{Q}_L^\sigma u = \sum_{l=0}^{\lceil L/\sigma \rceil} Q_l \mathcal{F}(u_{h_{\lceil \sigma L - \sigma^2 l \rceil}}) - \sum_{l=1}^{\lceil L/\sigma \rceil} Q_{l-1} \mathcal{F}(u_{h_{\lceil \sigma L - \sigma^2 l \rceil}}), \quad (35)$$

while for  $\sigma < 1$  it is more convenient to apply the version

$$\mathbf{Q}_L^\sigma u = \sum_{l=0}^{\lceil L\sigma \rceil} Q_{\lceil \frac{L}{\sigma} - \frac{l}{\sigma^2} \rceil} \mathcal{F}(u_{h_l}) - \sum_{l=1}^{\lceil L\sigma \rceil} Q_{\lceil \frac{L}{\sigma} - \frac{l}{\sigma^2} \rceil} \mathcal{F}(u_{h_{l-1}}). \quad (36)$$

Next, we state estimates on the convergence rate and computational cost of the anisotropic sparse grid  $\mathbf{Q}_L^\sigma$ . Since the proofs of these statements follow by the same arguments as in [5], most of the proofs will be omitted. Parts of the ideas will appear in the proofs in Section 5.2. First, we start with the cost for computing  $\mathbf{Q}_L^\sigma$  which is the analogon to Theorem 4.1 in [5].

**Theorem 5.** *The computational cost for the anisotropic sparse grid quadrature*

$$\mathbf{Q}_L^\sigma u = \sum_{\sigma l_1 + l_2 / \sigma \leq L} \left( \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} \right) u$$

measured in the number of quadrature points and degrees of freedom of the finite element method is

$$\text{cost}(\mathbf{Q}_L^\sigma u) \lesssim \begin{cases} 2^{L \max\{n_1/\sigma, n_2\sigma\}} & \text{if } n_1/\sigma \neq n_2\sigma, \\ 2^{Ln_2\sigma} L & \text{if } n_1/\sigma = n_2\sigma. \end{cases}$$

Similar to the regular sparse grid, the cost of the anisotropic version is essentially given by the cost needed for  $Q_{L/\sigma}^{(1)}$  or  $u_{h_{L\sigma}}$ . This is again a reduction compared to the full combination of finite element and quadrature method with  $\mathcal{O}(2^{L(n_1/\sigma + n_2\sigma)})$ . Furthermore, in the case of unequal computational cost ( $n_1/\sigma \neq n_2\sigma$ ), the method with the higher cost dominates the estimate.

Although the cost is reduced compared to the full grid, the convergence rate is nearly preserved, which is stated in the following theorem.

**Theorem 6.** *The anisotropic sparse grid*

$$\mathbf{Q}_L^\sigma u = \sum_{\sigma l_1 + l_2 / \sigma \leq L} \left( \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} \right) u$$

satisfies for the computation of the expectation  $\mathbb{E}[u(\mathbf{x})]$  ( $p = 1$ )

$$\|\mathbb{E}[u] - \mathbf{Q}_L^\sigma u\|_{H^1(D)} \lesssim \begin{cases} 2^{-L \min\{s_1/\sigma, s_2\sigma\}} \|f\|_{L^2(D)} & \text{if } s_1/\sigma \neq s_2\sigma, \\ 2^{-Ls_2\sigma} \sqrt{L} \|f\|_{L^2(D)} & \text{if } s_1/\sigma = s_2\sigma, \end{cases}$$

and for the second moment ( $p = 2$ )

$$\|\mathbb{E}[u^2] - \mathbf{Q}_L^\sigma u\|_{W^{1,1}(D)} \lesssim \begin{cases} 2^{-L \min\{s_1/\sigma, s_2\sigma\}} \|f\|_{L^2(D)}^2 & \text{if } s_1/\sigma \neq s_2\sigma, \\ 2^{-Ls_2\sigma} L \|f\|_{L^2(D)}^2 & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

*Proof.* For  $p = 2$ , we have by the mixed regularity (33)

$$\begin{aligned} \|\mathbb{E}[u] - \mathbf{Q}_L^\sigma u\|_{W^{1,1}(D)} &\leq \sum_{l_1\sigma + l_2/\sigma > L} \left\| \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} u \right\|_{W^{1,1}(D)} \\ &\lesssim \sum_{l_1\sigma + l_2/\sigma > L} 2^{-(s_1 l_1 + s_2 l_2)} \|f\|_{L^2(D)}^2 \\ &\lesssim \begin{cases} 2^{-L \min\{s_1/\sigma, s_2\sigma\}} \|f\|_{L^2(D)}^2 & \text{if } s_1/\sigma \neq s_2\sigma, \\ 2^{-Ls_2\sigma} L \|f\|_{L^2(D)}^2 & \text{if } s_1/\sigma = s_2\sigma, \end{cases} \end{aligned}$$

where the last estimate follows in the same way as in Theorem 4.3 of [5].

In the case  $p = 1$ , a stricter estimate can be shown by looking at the square of the  $H^1(D)$ -norm.

$$\begin{aligned} \|\mathbb{E}[u] - \mathbf{Q}_L^\sigma u\|_{H^1(D)}^2 &\lesssim \sum_{l_1\sigma + l_2/\sigma > L} \left\| \Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(2)} u \right\|_{H^1(D)}^2 \\ &\lesssim \sum_{l_1\sigma + l_2/\sigma > L} 2^{-2(s_1 l_1 + s_2 l_2)} \|f\|_{L^2(D)}^2, \end{aligned}$$

for which the same estimates hold such that we end up with

$$\|\mathbb{E}[u] - \mathbf{Q}_L^\sigma u\|_{H^1(D)}^2 \lesssim \begin{cases} 2^{-2L \min\{s_1/\sigma, s_2\sigma\}} \|f\|_{L^2(D)}^2 & \text{if } s_1/\sigma \neq s_2\sigma, \\ 2^{-2L s_2\sigma} L \|f\|_{L^2(D)}^2 & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

□

Hence, the smaller rate of the two numerical methods will determine the convergence rate of the anisotropic sparse grid method. The rate is only deteriorated if  $s_1/\sigma = s_2\sigma$ . Furthermore, we note that this estimate with  $\sigma = 1$  and  $s_1 = s_2$  is the setting considered in Theorem 4. Although the statement coincides for  $p = 2$ , we obtain an even better rate for the computation of the expectation, as the factor  $L$  is reduced to  $\sqrt{L}$ .

Combining both results leads to an error estimate in terms of the computational cost.

**Corollary 1.** *Let the mixed regularity assumption (33) hold and denote with  $N$  the computational cost for  $\mathbf{Q}_L^\sigma$ . Define*

$$\theta = \frac{\min\{s_1/\sigma, s_2\sigma\}}{\max\{n_1/\sigma, n_2\sigma\}}.$$

*Then, in the case of  $n_1/\sigma \neq n_2\sigma$ , the following convergence rate holds for  $p = 1, 2$ .*

$$\|\mathbb{E}[u^p] - \mathbf{Q}_L^\sigma u\|_X \lesssim \begin{cases} N^{-\theta} \|f\|_{L^2(D)}^p & \text{if } s_1/\sigma \neq s_2\sigma, \\ N^{-\theta} (\log N)^{p/2} \|f\|_{L^2(D)}^p & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

*If  $n_1/\sigma = n_2\sigma$ , we have*

$$\|\mathbb{E}[u^p] - \mathbf{Q}_L^\sigma u\|_X \lesssim \begin{cases} N^{-\theta} (\log N)^\theta \|f\|_{L^2(D)}^p & \text{if } s_1/\sigma \neq s_2\sigma, \\ N^{-\theta} (\log N)^{\theta+p/2} \|f\|_{L^2(D)}^p & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

*Furthermore, in the case  $\frac{s_1}{n_1} \neq \frac{s_2}{n_2}$  the optimal convergence rate  $\theta^* = \min\left\{\frac{s_1}{n_1}, \frac{s_2}{n_2}\right\}$  is achieved for all*

$$\sigma \in \left( \min \left\{ \sqrt{\frac{s_1}{s_2}}, \sqrt{\frac{n_1}{n_2}} \right\}, \max \left\{ \sqrt{\frac{s_1}{s_2}}, \sqrt{\frac{n_1}{n_2}} \right\} \right).$$

The corollary shows that the error of the sparse grid combining finite element method and quadrature, converges with the smaller of the two rates if the parameter  $\sigma$  is chosen in an optimal way. Only if the methods have the same approximation properties such that  $s_1/n_1 = s_2/n_2$ , is the rate reduced by a logarithmic term.

### 5.1.1 Numerical Examples

For an illustration of the anisotropic combination technique, we consider again the constructed example with one stochastic and one spatial variable, where the input of the PDE is given by (30).

While we applied a regular sparse grid, which corresponds to the choice  $\sigma = 1$  with equilibrated approximation powers  $s_1 = s_2$ , in Section 4, we want to illustrate here that using the anisotropic version we can achieve a better convergence rate, as no logarithmic factor deteriorates the convergence rate.

Only for the case where we considered the  $L^2(D)$  error and applied the trapezoidal rule for the quadrature, the optimal convergence rate was achieved, because in this case  $\frac{s_1}{n_1} = \frac{s_2}{n_2}$  and the accuracy and cost of the methods are balanced with the same parameter  $\sigma$ .

Furthermore, we mention that the obtained convergence rates in the last section coincide with the predicted rates in Theorem 6 and Corollary 1. The observed smaller exponents of the logarithmic terms are consistent with the improvement in Theorem 6 for  $p = 1$ .

Here, we exemplify the convergence of the anisotropic sparse grid for the computation of the expectation with respect to the  $H^1(D)$ -norm for different  $\sigma$ . For the computation of the second moment, we provide only the results for the optimal choice of  $\sigma$ , as the convergence behavior is essentially the same. In both cases, we choose  $N_l = 2^l$  points for the mesh in  $D$  and for the quadrature method. Hence,  $n_1 = n_2 = 1$ . We use the approximation power of the numerical methods to choose the values of  $\sigma$ . The expected rates of the anisotropic sparse grid in terms of the computational cost are shown for three choices in Table 1.

	$\sigma^2 = \frac{n_1}{n_2} = 1$	$\sigma^2 = \frac{s_1+n_1}{s_2+n_2}$	$\sigma^2 = \frac{s_1}{s_2}$
Monte Carlo	$N^{-\frac{1}{2}} \log(N)^{\frac{1}{2}}$	$N^{-\frac{1}{2}}$	$N^{-\frac{1}{2}} \log(N)^{\frac{1}{2}}$
Trapezoidal	$N^{-1} \log(N)$	$N^{-1}$	$N^{-1} \log(N)^{\frac{1}{2}}$
Gauss-Legendre	$N^{-1} \log(N)$	$N^{-1}$	$\left(N^{-1} \log(N)^{\frac{1}{2}}\right)$

Table 1: Predicted rates of convergence.

The convergence results are presented in Figure 7. The curves show an erratic behavior because the probabilistic nature of the Monte Carlo method only provides an error rate in the mean square sense. Furthermore, the implementation of the anisotropic method includes indices with  $l \leq \lceil L/\sigma \rceil$  and  $l \leq \lceil L\sigma \rceil$ , so that the increase in levels considered in each direction is not equal for growing sparse grid levels  $L$ .

The upper left panel depicts the convergence rates for the trapezoidal rule. From Corollary 1 we expect that the optimal convergence rate of 1 is achieved for

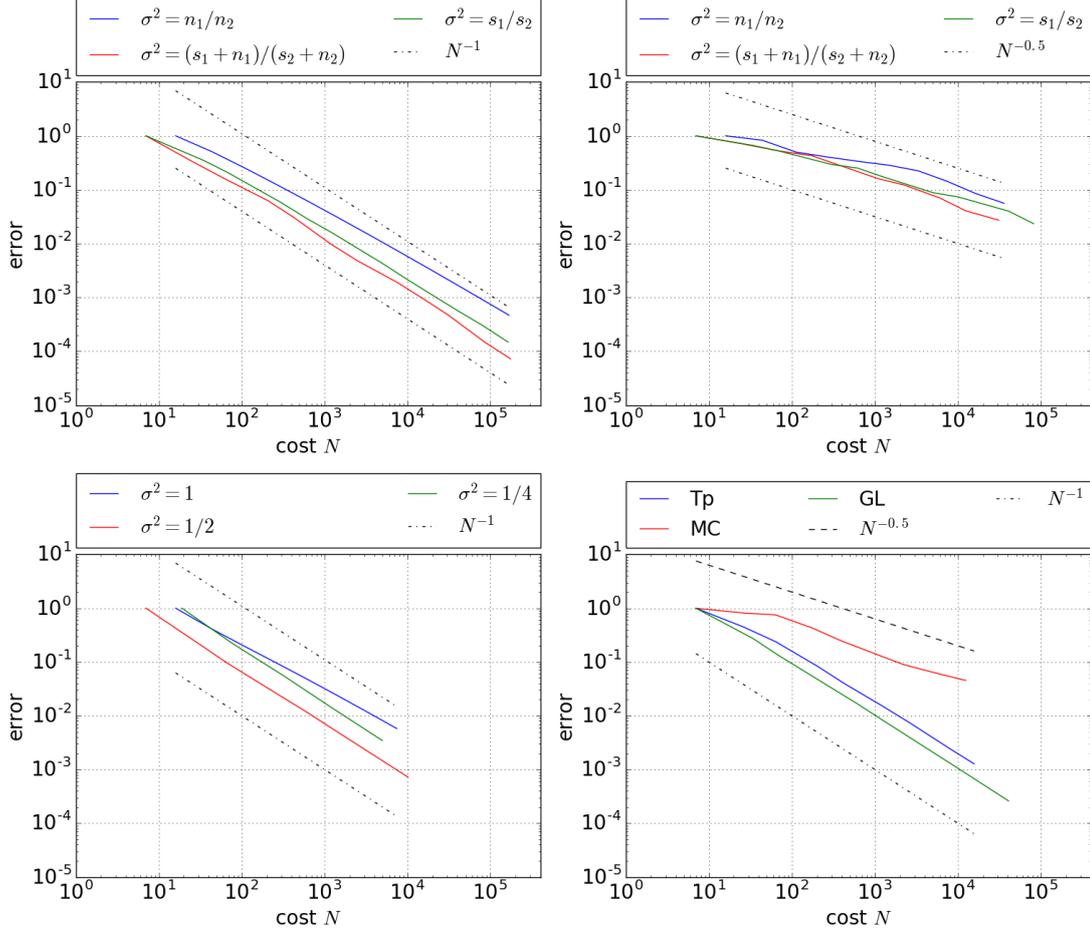


Figure 7: Convergence of anisotropic sparse grid method for the computation of the expected value measured in  $H^1(D)$ -norm with different quadrature rules. Top left: trapezoidal rule, top right: Monte Carlo method, bottom left: Gauss-Legendre. Bottom right: convergence result for the computation of the second moment with optimal choice of  $\sigma$ .

$\sigma^2 \in (\frac{1}{2}, 1)$ . Indeed, the numerical implementation exhibits this convergence rate for the choice of an equilibrated benefit-cost ratio. For  $\sigma^2$  chosen as the endpoint of this interval, the convergence is, however, affected by a logarithmic term, which results in slightly slower convergence.

Using the Monte Carlo quadrature (upper right panel) instead provides similar results with an optimal interval  $\sigma^2 \in (1, 2)$ . However, the rate of the finite element method is not achieved, because the Monte Carlo only converges with a rate  $1/2$ . The Gauss-Legendre quadrature instead has a higher convergence rate, as it converges exponentially for analytic functions. Hence, fewer levels should be used for the  $y$ -direction than for the  $x$ -direction. For all  $s_2 \in \mathbb{N}$ , it holds  $u \in C^{s_2}$  in the stochastic variable such that the Gauss-Legendre quadrature satisfies the asymptotic convergence rate  $\mathcal{O}(2^{-s_2})$ . Hence, any choice for  $\sigma$  with  $\sigma^2 \in (0, 1)$  will achieve the convergence of order 1. This can be observed in the lower left panel which provides the convergence analysis for  $\sigma^2 = \frac{1}{4}, \frac{1}{2}, 1$ . Only for  $\sigma = 1$  the rate is lowered, as in this case the cost is equilibrated.

Overall, we observe that the convergence rate of the anisotropic sparse grid is limited by the lower of the two rates of the separate numerical approximations. Hence, if the regularity is provided, the Gaussian and trapezoidal quadrature rules yield the faster convergence. Furthermore, it can be achieved that the lower rate of both methods is not deteriorated by a logarithmic factor if the approximation power and cost are not equal for both methods, which is an improvement compared to applying a regular sparse grid.

Finally, we observe in the lower right panel of Figure 7 that this also holds for the computation of the second moment.

Although we presented here only numerical examples of the anisotropic sparse grid for the simple case of one-dimensional spatial and stochastic domain, the theoretical results also hold for higher-dimensional quadrature rules which satisfy the condition (33). For example, a sparse grid quadrature in  $\mathbf{y}$  might be considered, which exhibit an convergence of  $\mathcal{O}(2^{-rl}l^{m-1})$  using  $2^l l^{m-1}$  quadrature points. By slightly reducing the rate  $r$ , this quadrature fits the considered setting, as  $l^t \lesssim 2^{l\varepsilon}$  for  $\varepsilon > 0$  and  $t > 0$ . If the the stochastic variables are not equally important, one could also consider an anisotropic quadrature method in the multidimensional case, where the levels for the quadrature are coupled to the first random variable, such that the error and also the cost is of the required form.

## 5.2 Sparse Grid for Truncation and Quadrature

After having addressed a sparse grid within the stochastic variables and in combination with the PDE solver, we consider in this section a different part of the computation of the quantities of interest. When the random field is written in the Karhunen-Loeve expansion and subsequently parametrized, the exact solution of the PDE depends on infinitely many variables. The computation of the expectation or of a higher moment of the partial differential solution hence corresponds to an integral over an infinite-dimensional parameter space.

So far, the truncation to a finite-dimensional parameter space has been fixed. In this section, we will now look at the interplay between the finite noise truncation and stochastic approximation, while disregarding the spatial discretization. We will use a sparse grid approach to reduce the complexity of computing a quantity which requires a truncation of an infinite-dimensional stochastic.

To that end, we consider a simpler problem. Instead of looking at a PDE solution, we compute the expectation of a given function that depends on infinitely many random variables  $\mathbf{y} = (y_1, y_2, \dots)$  as a quantity of interest.

For  $i \in \mathbb{N}$ , let  $y_i$  be independent uniformly distributed random variables taking values in  $\Gamma = [-\frac{1}{2}, \frac{1}{2}]$  and define the infinite-dimensional parameter space  $\mathbf{\Gamma} = \times_{k=1}^{\infty} \Gamma_k$ . Then, for a function  $G: \mathbf{\Gamma} \rightarrow \mathbb{R}$ , we are interested in the value

$$\mathbb{E}_{\mathbf{y}}[G(\mathbf{y})] = \int_{\mathbf{\Gamma}} G(\mathbf{y})\rho(\mathbf{y})d\mathbf{y} = \int_{\mathbf{\Gamma}} G(\mathbf{y})d\mathbf{y}. \quad (37)$$

with  $\rho(\mathbf{y}) \equiv 1$ . A rigorous definition of this integral is more delicate, as  $d\mathbf{y}$  is

a measure on an infinite-dimensional space. In infinite dimensions, there is no analogue to the Lebesgue measure ([55]). However, it is possible to construct a product measure on  $\Gamma$  (see [56]). The integral is then defined as the limit

$$\mathbb{E}_{\mathbf{y}}[G(\mathbf{y})] = \lim_{m \rightarrow \infty} \int_{\Gamma^m} G(\mathbf{y}_{1:m}, 0, \dots) d\mathbf{y}_{1:m},$$

where we use the notation  $\mathbf{y}_{1:m} = (y_1, \dots, y_m)$ .

The existence of this limit can be shown under certain conditions on the function  $G$ . In [45], the condition of  $G$  being in a certain reproducing kernel Hilbert space was used. The existence of the limit requires the continuity of  $G$  and a certain decrease of influence of the variables  $y_k$  on the value  $G(\mathbf{y})$  as  $k \rightarrow \infty$ . Hereinafter, we assume that the integral is well-defined. If the variables derive from a Karhunen-Loeve expansion, an unequal dependence on the stochastic variables is given, because as stated in Section 3.3 the coefficients in the Karhunen-Loeve expansion exhibit certain decay properties depending on the regularity of the field. By Lemma 5 this anisotropy is also maintained in the PDE solution.

An approximation of (37) requires first a reduction of  $G$  to a function  $G_m : \Gamma^m \rightarrow \mathbb{R}$  on a finite-dimensional domain and subsequently a replacement of the integral by a quadrature rule. For the first step, we define the  $m$ -dimensional surrogate  $G_m$  by

$$G_m(\mathbf{y}_{1:m}) = G(y_1, \dots, y_m, 0, 0, \dots).$$

Only the first  $m$  variables  $\mathbf{y}_{1:m} = (y_1, \dots, y_m)$  are considered, while for  $k > m$  the function is evaluated at a fixed anchor point  $\bar{y} = 0$ . This choice is natural for the centered variables  $y_k$ . Viewed from another perspective, this means the function  $G$  is evaluated at the projection of  $\mathbf{y}$  onto the finite-dimensional space

$$\{\mathbf{y} \in \mathbb{R}^{\mathbb{N}} : y_i = 0 \text{ for } i > m\}.$$

Thus, we can obtain an approximation of the integral by  $\mathbf{Q}G_m$  with  $\mathbf{Q}$  being an  $m$ -dimensional quadrature method.

Since we want to use a sparse grid effect between the projection onto a finite-dimensional parameter space and the quadrature rule, we need a sequence of truncation levels and quadrature methods. To that end, let  $\{\mathbf{Q}_l\}_l$  be a sequence of quadrature rules with increasing accuracy. For  $\mathbf{Q}_l$ , we assume for any finite-dimensional function  $v : \mathbb{R}^m \rightarrow \mathbb{R}$  the error bounds

$$|(\mathbf{I} - \mathbf{Q}_l)v| \lesssim 2^{-s_1 l} \|v\|_{\mathcal{Y}} \quad (38)$$

with a constant that is independent of  $m$  and where the norm  $\|\cdot\|_{\mathcal{Y}}$  depends on the quadrature rule. We remark that the assumption includes that the quadrature  $\mathbf{Q}_l$  can be applied to any finite-dimensional function and that the estimate is independent of  $m$ .

Furthermore, we assume that  $\lim_{l \rightarrow \infty} \mathbf{Q}_l v = \int v(\mathbf{y}_{1:m}) d\mathbf{y}_{1:m}$  such that

$$\begin{aligned} \int v(\mathbf{y}_{1:m}) d\mathbf{y}_{1:m} &= \sum_{l=0}^{\infty} (\mathbf{Q}_l v - \mathbf{Q}_{l-1} v) + \mathbf{Q}_0 v \\ &= \sum_{l=0}^{\infty} \Delta_l^Q v \end{aligned} \quad (39)$$

with the difference operator defined as in (31).

In the same manner, we consider an increasing sequence of truncation levels

$$m_0 < m_1 \leq m_2 \leq \dots \leq m_l \leq \dots < \infty$$

such that the function evaluation can be written as

$$\begin{aligned} G(\mathbf{y}) &= \sum_{l=1}^{\infty} \left( G(y_1, \dots, y_{m_l}, 0, \dots) - G(y_1, \dots, y_{m_{l-1}}, 0, \dots) \right) + G(\mathbf{y}_{1:m_0}, 0, \dots) \\ &= \sum_{l=0}^{\infty} \Delta_l G(\mathbf{y}) \end{aligned} \quad (40)$$

with

$$\Delta_l G(\mathbf{y}) = \begin{cases} G_{m_l}(\mathbf{y}_{1:m_l}) - G_{m_{l-1}}(\mathbf{y}_{1:m_{l-1}}) & \text{if } l \geq 1 \\ G_{m_1}(\mathbf{y}_{1:m_0}) & \text{if } l = 0. \end{cases}$$

We assume that the truncation levels  $\{m_l\}_l$  are chosen such that the error made by the truncation decreases according to

$$\|G(\mathbf{y}) - G_{m_l}(\mathbf{y}_{1:m_l})\|_{\mathcal{Y}} \lesssim 2^{-s_2 l} \quad (41)$$

for some  $s_2 > 0$ .

The anisotropic sparse grid integral approximation then takes the form

$$\mathbf{Q}_L^\sigma G := \sum_{\sigma l_1 + l_2 / \sigma \leq L} \Delta_{l_1}^Q \Delta_{l_2} G. \quad (42)$$

This is equivalent to the formulations (cf. (35) and (36))

$$\mathbf{Q}_L^\sigma G = \sum_{l=0}^{L/\sigma} \mathbf{Q}_l G_{\sigma L - \sigma^2 l} - \sum_{l=1}^{L/\sigma} \mathbf{Q}_{l-1} G_{\sigma L - \sigma^2 l}, \quad (43)$$

respectively

$$\mathbf{Q}_L^\sigma G = \sum_{l=0}^{L\sigma} \mathbf{Q}_{\frac{l}{\sigma} - \frac{l}{\sigma^2}} G_l \sum_{l=1}^{L\sigma} \mathbf{Q}_{\frac{l}{\sigma} - \frac{l}{\sigma^2}} G_{l-1}. \quad (44)$$

We note that in (43) and (44) the quadrature rules are applied to functions on different finite-dimensional spaces as the truncation level of  $G$  changes. Hence, the sequence of quadrature rules must be capable of approximating the integral over any finite-dimensional space. For the analysis of the error of the sparse grid formulation, it is important that the error bounds and cost estimations for  $\mathbf{Q}_l$  are independent of the dimension of the underlying space.

The straightforward choice for the quadrature is a Monte Carlo or Quasi Monte Carlo quadrature rule, as their convergence rates only depend on the number of quadrature points and not on the dimension. We will consider the Monte Carlo quadrature in Section 5.2.1. As the convergence of the Monte Carlo method is rather slow, a product of one-dimensional quadrature rules is considered in the second part. There the decrease of importance of the variables is used to construct a sequence of product rules that have error and cost bounds independent of the dimension.

Before we look at these different quadrature methods in detail, we can prove an error bound for the representation (42).

**Theorem 7.** *Let the conditions (38) and (41) hold. Then the approximation  $\mathbf{Q}_L^\sigma G$  defined by (42) satisfies*

$$\left| \int_{\Gamma} G(\mathbf{y}) d\mathbf{y} - \mathbf{Q}_L^\sigma G \right| \lesssim \begin{cases} 2^{-L \min\{s_1/\sigma, s_2\sigma\}} & \text{if } s_1/\sigma \neq s_2\sigma \\ 2^{-L s_1/\sigma} L & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

*Proof.* Using the expansion (39) and (40), we get

$$\begin{aligned} |\mathbb{E}_{\mathbf{y}} [G(\mathbf{y})] - \mathbf{Q}_L^\sigma G| &= \left| \int_{\Gamma} G(\mathbf{y}) d\mathbf{y} - \sum_{l_1\sigma + l_2/\sigma \leq L} \Delta_{l_1}^Q \Delta_{l_2} G \right| \\ &\leq \sum_{l_1\sigma + l_2/\sigma > L} |\Delta_{l_1}^Q \Delta_{l_2} G| \leq \sum_{l_1\sigma + l_2/\sigma > L} \left\| \Delta_{l_1}^Q \right\|_{\mathbf{y} \rightarrow \mathbb{R}} \|\Delta_{l_2} G\|_{\mathbf{y}}. \end{aligned}$$

The assumptions (41) and (38) imply

$$\begin{aligned} |\Delta_{l_1}^Q v| &\leq |(\mathbf{I} - \mathbf{Q}_{l_1})v| + |(\mathbf{I} - \mathbf{Q}_{l_1-1})v| \\ &\leq 2^{-s_1 l_1} \|v\|_{\mathbf{y}} + 2^{-s_1(l_1-1)} \|v\|_{\mathbf{y}} \lesssim 2^{-s_1 l_1} \|v\|_{\mathbf{y}} \end{aligned}$$

and analogously,

$$\|\Delta_{l_2} G\|_{\mathbf{y}} \leq \|G - G_{m_{l_2}}\|_{\mathbf{y}} + \|G - G_{m_{l_2-1}}\|_{\mathbf{y}} \lesssim 2^{-s_2 l_2}.$$

Therefore,

$$\begin{aligned}
|\mathbb{E}_{\mathbf{y}} [G(\mathbf{y})] - \mathbf{Q}_L^\sigma G| &\lesssim \sum_{l_1\sigma + l_2/\sigma > L} 2^{-s_1 l_1} 2^{-s_2 l_2} \\
&\leq \sum_{l_1=0}^{L/\sigma} \sum_{l_2=L\sigma - l_1\sigma^2 + 1}^{\infty} 2^{-s_1 l_1} 2^{-s_2 l_2} + \sum_{l_1=L/\sigma + 1}^{\infty} \sum_{l_2=0}^{\infty} 2^{-s_1 l_1} 2^{-s_2 l_2} \\
&\lesssim \sum_{l_1=0}^{L/\sigma} 2^{-s_1 l_1 - s_2\sigma L + s_2 l_1 \sigma^2} + \sum_{l_1=L/\sigma + 1}^{\infty} 2^{-s_1 l_1} \\
&\lesssim 2^{-s_2\sigma L} \sum_{l_1=0}^{L/\sigma} 2^{-l_1(s_1 - s_2\sigma^2)} + 2^{-s_1 L/\sigma} \\
&= 2^{-s_2\sigma L} \left( \sum_{l_1=0}^{L/\sigma} 2^{-l_1\sigma(s_1/\sigma - s_2\sigma)} + 2^{-L(s_1/\sigma - s_2\sigma)} \right).
\end{aligned}$$

We distinguish between three cases. If  $s_1/\sigma > s_2\sigma$  the exponents are negative such that

$$|\mathbb{E}_{\mathbf{y}} [G(\mathbf{y})] - \mathbf{Q}_L^\sigma G| \lesssim 2^{-s_2\sigma L} (1 + 1) \lesssim 2^{-s_2\sigma L}.$$

If instead  $s_1\sigma < s_2\sigma$ , we find

$$|\mathbb{E}_{\mathbf{y}} [G(\mathbf{y})] - \mathbf{Q}_L^\sigma G| \lesssim 2^{-s_2\sigma L} (2^{-L(s_1/\sigma - s_2\sigma)} + 2^{-L(s_1/\sigma - s_2\sigma)}) \lesssim 2^{-s_1\sigma L}.$$

For  $s_1\sigma = s_2\sigma$ , we have

$$|\mathbb{E}_{\mathbf{y}} [G(\mathbf{y})] - \mathbf{Q}_L^\sigma G| \lesssim 2^{-s_2\sigma L} \left( \sum_{l_1=0}^{L/\sigma} 1 + 1 \right) \lesssim 2^{-s_2\sigma L} \frac{L}{\sigma}.$$

□

This bound holds, due to the equivalence of the formulations, also for the approximations obtained by the combination technique (43) or (44). The theorem is similar to Theorem 6, but here the quadrature is combined with different truncation levels instead of finite element solutions. The theorem shows that the rate of convergence is essentially given by the smaller of the quadrature and truncation convergence rates given by condition (38) and (41). Since  $\{G_{m_l}\}_l$  depends on a varying number of variables which is unbounded, the evaluation is not constant anymore. A cost model for  $\mathbf{Q}_{l_1} G_{m_{l_2}}$ , therefore, needs to account for this. For example, the cost of  $\mathbf{Q}_{l_1} G_{m_{l_2}}$  can be measured in terms of the number of quadrature points in  $\mathbf{Q}_{l_1}$  times the number of variables  $m_{l_2}$  that are considered in the evaluation of  $G$ . Consequently, the cost of  $\mathbf{Q}_L^\sigma G$  is reduced compared to  $\mathbf{Q}_{L/\sigma} G_{m_{L\sigma}}$ , while up to a logarithmic term the approximation power is the same. The exact cost of computing  $\mathbf{Q}_L^\sigma G$  depends on the particular choice of the sequence  $\{m_l\}_l$  and the cost of the quadrature rules.

### 5.2.1 Monte Carlo

In this section, we examine how a Monte Carlo method can be applied in combination with the truncation level. For the Monte Carlo quadrature to be applicable, we need that  $G$  is square integrable, since then  $G_m$  is square integrable for any  $m$ . Let  $\mathbf{Q}_l$  be the Monte Carlo quadrature with  $N_l = 2^l$  sample points. The estimate (38) is then satisfied in the root mean square error sense

$$\sqrt{\mathbb{E} [ |(\mathbf{I} - \mathbf{Q}_l)v|^2 ]} \lesssim 2^{-\frac{1}{2}l} \|v\|_{L^2_\rho(\Gamma^m)}.$$

Hence, if the condition (cf. (41))

$$\|G(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{L^2_\rho(\mathbf{\Gamma})} \lesssim 2^{-s_2 l} \quad (45)$$

is satisfied, the statement of Theorem 7 holds in the root mean square sense.

The condition (45) measures how much of the value  $G$  is covered by the first  $m_l$  variables. To quantify this, we consider a certain class of functions. Note that this assumption can be weakened, for example to Hölder-continuous functions.

**Assumption 1.** *Assume  $G: \mathbf{\Gamma} \rightarrow \mathbb{R}$  is continuously differentiable and for every  $k \in \mathbb{N}$*

$$\sup_{\mathbf{y} \in \mathbf{\Gamma}} |\partial_{y_k} G(\mathbf{y})| \lesssim \gamma_k$$

for a decreasing sequence  $\{\gamma_k\}_k$  that is summable.

First of all, the summability of  $\{\gamma_k\}_k$  ensures that  $G$  is square integrable. Here two choices for the sequence  $\{\gamma_k\}_k$  are considered. Either the sequence decays algebraically, meaning

$$\gamma_k \lesssim k^{-\alpha}$$

with  $\alpha > 1$  such that  $\{\gamma_k\}_k \in \ell^1$ , or with an exponential rate

$$\gamma_k \lesssim 2^{-\beta k}$$

for some  $\beta > 0$ . These choices are motivated by the decay rates for the coefficients in the Karhunen-Loeve expansion of a random field. If  $a(\omega, x) \in L^2(\Omega, H^s(D))$ , then the coefficients satisfy an algebraic decay rate, while for an analytic covariance function for the random field the decay is of the second type (cf. Section 3.3).

Under the above assumption, we can show that condition (45) is satisfied when the levels  $m_l$  are chosen appropriately.

**Lemma 6.** *Let  $G: \mathbf{\Gamma} \rightarrow \mathbb{R}$  satisfy Assumption 1.*

(i) *If  $\gamma_k \lesssim k^{-\alpha}$  for  $\alpha > 1$  and if the truncation levels are chosen according to  $m_l = 2^l$ , then*

$$\|G(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{L^2_\rho(\mathbf{\Gamma})} \lesssim 2^{-(\alpha-1)l}.$$

(ii) If instead  $\gamma_k \lesssim 2^{-\beta k}$  for  $\beta > 0$  and  $m_l = l$ , then

$$\|G(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{L^2_\rho(\Gamma)} \lesssim 2^{-\beta l}.$$

*Proof.* (i) Using the mean value theorem, we have, for  $M > m_l$ ,

$$\begin{aligned} & \|G_M(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{L^2_\rho(\Gamma)} \leq \sup_{\mathbf{y} \in \Gamma} |G(\mathbf{y}_{1:M}) - G_{m_l}(\mathbf{y}_{1:M})| \\ & \leq \sum_{k=m_l+1}^M \sup_{\mathbf{y} \in \Gamma} |\partial_{y_k} G(\mathbf{y})| \sup_{y_k \in \Gamma} |y_k| \\ & \lesssim \sum_{k=m_l+1}^M \gamma_k. \end{aligned}$$

Then taking the limit  $M \rightarrow \infty$  yields together with the lower semi-continuity of the norm

$$\begin{aligned} \|G(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{L^2_\rho(\Gamma)} & \lesssim \sum_{k=m_l+1}^{\infty} \gamma_k \lesssim \sum_{k=2^l+1}^{\infty} k^{-\alpha} = \sum_{i=l}^{\infty} \sum_{k=2^i+1}^{2^{i+1}} k^{-\alpha} \\ & \lesssim \sum_{i=l}^{\infty} 2^i 2^{-\alpha i} \lesssim 2^{(1-\alpha)l}, \end{aligned}$$

where in the last step it was used that  $\alpha > 1$ .

(ii) Similarly,

$$\|G(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{L^2_\rho(\Gamma)} \lesssim \sum_{k=m_l+1}^{\infty} \gamma_k \lesssim \sum_{k=l+1}^{\infty} 2^{-\beta k} \lesssim 2^{-\beta l}.$$

□

We now turn to the cost estimation of the sparse grid formulation computed by the combination technique formula (43) or (44). We provide an upper bound for the cost for the two different choices of truncation levels.

**Theorem 8.** *The cost of computing  $\mathbf{Q}_L^\sigma G$  is given*

(i) for the truncation levels  $m_l = 2^l$  by

$$\text{cost}(\mathbf{Q}_L^\sigma G) \lesssim \begin{cases} 2^{L \max\{1/\sigma, \sigma\}} & \text{if } \sigma \neq 1, \\ 2^{\sigma L} L & \text{if } \sigma = 1, \end{cases}$$

(ii) and for the choice  $m_l = l$  by

$$\text{cost}(\mathbf{Q}_L^\sigma G) \lesssim 2^{L/\sigma}.$$

*Proof.* (i) We recall that the number of the quadrature points used for  $\mathbf{Q}_l$  was chosen as  $N_l = 2^l$ . We assume w.l.o.g. that  $\mathbf{Q}_L^\sigma G$  is computed according to (43).

$$\begin{aligned} \text{cost}(\mathbf{Q}_L^\sigma G) &= \sum_{l=0}^{L/\sigma} \text{cost}(\mathbf{Q}_l G_{\sigma L - \sigma^2 l}) + \sum_{l=1}^{L/\sigma} \text{cost}(\mathbf{Q}_{l-1} G_{\sigma L - \sigma^2 l}) \\ &\leq \sum_{l=0}^{L/\sigma} 2^l 2^{(\sigma L - \sigma^2 l)} + \sum_{l=1}^{L/\sigma} 2^{(l-1)} 2^{(\sigma L - \sigma^2 l)} \\ &\lesssim \sum_{l=0}^{L/\sigma} 2^l 2^{(\sigma L - \sigma^2 l)} = 2^{\sigma L} \sum_{l=0}^{L/\sigma} 2^{l(1-\sigma^2)}. \end{aligned}$$

For  $\sigma > 1$ , we obtain  $\text{cost}(\mathbf{Q}_L^\sigma G) \lesssim 2^{L\sigma}$ , while for  $\sigma < 1$  the estimate  $\text{cost}(\mathbf{Q}_L^\sigma G) \lesssim 2^{\sigma L} 2^{L/\sigma(1-\sigma^2)} = 2^{L/\sigma}$  holds. In the case  $\sigma^2 = 1$  we end up with  $\text{cost}(\mathbf{Q}_L^\sigma G) \lesssim 2^{\sigma L} \sum_{l=0}^{L/\sigma} 1 \lesssim 2^{\sigma L} L$ .

(ii) Similarly,

$$\begin{aligned} \text{cost}(\mathbf{Q}_L^\sigma G) &= \sum_{l=0}^{L/\sigma} \text{cost}(\mathbf{Q}_l G_{\sigma L - \sigma^2 l}) + \sum_{l=1}^{L/\sigma} \text{cost}(\mathbf{Q}_{l-1} G_{\sigma L - \sigma^2 l}) \\ &\leq \sum_{l=0}^{L/\sigma} 2^l (\sigma L - \sigma^2 l) + \sum_{l=1}^{L/\sigma} 2^{(l-1)} (\sigma L - \sigma^2 l) \\ &\lesssim \sum_{l=0}^{L/\sigma} 2^l (\sigma L - \sigma^2 l) = \sigma^2 \sum_{l=0}^{L/\sigma} 2^{L/\sigma - l} \lesssim 2^{L/\sigma}. \end{aligned}$$

□

An error bound in terms of the cost is given by the following corollary.

**Corollary 2.** *Let  $G$  be like in Lemma 6,  $m_l = 2^l$ , and set*

$$\theta = \frac{\min\{\frac{1}{2\sigma}, (\alpha - 1)\sigma\}}{\max\{\frac{1}{\sigma}, \sigma\}}.$$

*Then*

$$\left| \int_{\Gamma} G(\mathbf{y}) d\mathbf{y} - \mathbf{Q}_L^\sigma G \right| \lesssim \begin{cases} N^{-\theta} & \text{if } \sigma \neq 1 \text{ and } \sigma^2 \neq \frac{1}{2(\alpha-1)}, \\ N^{-\theta} \log N & \text{if } \sigma^2 = \frac{1}{2(\alpha-1)} \text{ and } \alpha \neq \frac{3}{2} \\ N^{-\theta} (\log N)^\theta & \text{if } \sigma = 1 \text{ and } \alpha \neq \frac{3}{2} \\ N^{-\theta} (\log N)^{(1+\theta)} & \text{if } \sigma = 1 \text{ and } \alpha = \frac{3}{2}. \end{cases}$$

In the second case, set

$$\theta = \min \left\{ \frac{1}{2}, \beta\sigma^2 \right\}$$

and obtain

$$\left| \int_{\mathbf{\Gamma}} G(\mathbf{y}) d\mathbf{y} - \mathbf{Q}_L^\sigma G \right| \lesssim \begin{cases} N^{-\theta} & \text{if } \sigma^2 \neq \frac{1}{2\beta}, \\ N^{-\theta} \log N & \text{if } \sigma^2 = \frac{1}{2\beta}. \end{cases}$$

*Proof.* The statement follows from combining Lemma 6 and the Theorems 7 and 8.  $\square$

By maximizing  $\theta$  in Corollary 2, the optimal choice for  $\sigma$  can be found. In the first case, the optimal rate of  $\theta = \min\{\frac{1}{2}, \alpha - 1\}$  is obtained by

$$\sigma^2 \in \left( \min \left\{ 1, \frac{1}{2(\alpha - 1)} \right\}, \max \left\{ 1, \frac{1}{2(\alpha - 1)} \right\} \right).$$

Hence, the rate is limited either by the decay in the variables or the Monte Carlo method. In the second case, the importance of  $y_k$  decreases faster such that the optimal rate  $\theta = \frac{1}{2}$  is achieved for

$$\sigma^2 > \frac{1}{2\beta}.$$

### 5.2.2 Product Quadrature Rule

In this section, we will consider a quadrature rule that is a product of one-dimensional quadrature rules instead of a Monte Carlo quadrature. The motivation behind this is that, as seen above, using a Monte Carlo method, the convergence rate for the sparse grid can be at most of order 1/2, as the Monte Carlo method does not achieve a higher rate. An alternative might be to consider a product of one-dimensional quadrature rules with higher convergence rates.

However, unlike the Monte Carlo method, a tensor product built from one-dimensional quadrature rules is not independent of the dimension, as it suffers from the curse of dimensionality. Furthermore, for a combination with the truncation of the parameter space, we need to have a sequence of quadrature rules whose convergence rates do not depend on the dimension of the integration domain. Simply applying the product quadrature with the same number of points in each direction would therefore not lead to a useful quadrature rule, even if a sparse version is considered. However, we can use the knowledge that the variables  $y_i$  are not equally important to construct an anisotropic product quadrature method whose cost and error can be estimated independently of the dimension. The construction of a sequence of such quadrature methods is discussed next.

For that, we consider a sequence of quadrature rules  $\{\mathbf{Q}_l\}_l$  of the form

$$\mathbf{Q}_l = Q_{N_1^{(l)}}^{(1)} \otimes \dots \otimes Q_{N_m^{(l)}}^{(m)},$$

where the number of points  $N_k^{(l)}$  for each  $k = 1, \dots, m$  is determined by the structure of the integrand. Hence, only the level parameter  $l$  controls the accuracy of the quadrature. In the following, we will discuss this product quadrature where  $Q_{N_k^{(l)}}^{(k)}$  are one-dimensional Gauss-Legendre quadratures because, in the case of smooth integrands, they exhibit exponential convergence rates. Although we have shortly mentioned the product quadrature in Section 3.1, we will here consider the specific choice of Gaussian quadratures in more detail and construct a rule for possibly varying stochastic dimensions.

We emphasize that a full, but anisotropic, tensor product of quadrature rules is considered. Alternatively, one could also study an anisotropic sparse grid quadrature method, which, for example, has been studied for the stochastic collocation method in [19, 39] and in a multilevel formulation balancing the PDE solver and the quadrature method in [21, 23]. Here, however, we consider only the case of a full product.

Before looking at how to balance a Gauss-Legendre product quadrature with the truncation level, we need to assess the accuracy of the product quadrature. Following ideas in [15], the number of quadrature points in each direction can be chosen such that the product quadrature satisfies condition (38) independently of the dimension.

**Lemma 7.** *Let  $v \in C(\Gamma^m) \subset L_\rho^2(\Gamma^m)$  be a function, such that for all  $k = 1, \dots, m$ , it is analytically extendable into  $\Sigma(\Gamma_k, \tau_k) = \{z \in \mathbb{C} : \text{dist}(z, \Gamma_k) \leq \tau_k\}$ . If a tensor product of Gauss-Legendre quadratures with*

$$N_k^{(l)} \geq \frac{\ln(2)}{2h(\tau_k)}l + \frac{1}{2}$$

*quadrature points is considered, where  $h(\tau_k) = \ln(2\tau_k + \sqrt{1 + 4\tau_k^2})$ , then the error is bounded by*

$$|(\mathbf{I} - \mathbf{Q}_l)v| \lesssim 2^{-l} \max_{k=1, \dots, m} \|v\|_{C^0(\Sigma(\Gamma_k, \tau_k); C^0(\Gamma_k^*))} \sum_{k=1}^m g(\tau_k)$$

*with  $g(\tau) = \frac{4}{(2\tau + \sqrt{1 + 4\tau^2} - 1)}$  and  $\Gamma_k^*$  as in (7). If in addition  $\{g(\tau_k)\}_k$  is summable, the bound is independent of  $m$ :*

$$|(\mathbf{I} - \mathbf{Q}_l)v| \lesssim 2^{-l} \max_{k=1, \dots, m} \|v\|_{C^0(\Sigma(\Gamma_k, \tau_k); C^0(\Gamma_k^*))}.$$

*Proof.* The error of the  $m$ -dimensional quadrature rule can be estimated in terms of the one-dimensional errors as

$$|(\mathbf{I} - \mathbf{Q})v| \leq \sum_{k=1}^m |(Q^{(1)} \otimes \dots \otimes Q^{(k-1)} \otimes (I^{(k)} - Q^{(k)}) \otimes I^{(k+1)} \otimes \dots \otimes I^{(m)})v|.$$

As the quadrature  $Q : C^0(\Gamma) \rightarrow \mathbb{R}$  and the integral operator  $I : C^0(\Gamma) \rightarrow \mathbb{R}$  are continuous with constant 1, it holds that

$$\begin{aligned} & | (Q^{(1)} \otimes \dots \otimes Q^{(k-1)} \otimes (I^{(k)} - Q^{(k)}) \otimes I^{(k+1)} \otimes \dots \otimes I^{(m)}) v | \\ & \lesssim \sup_{\mathbf{y}_k^* \in \Gamma_k^*} | (I^{(k)} - Q^{(k)}) v(\mathbf{y}_k^*) |, \end{aligned}$$

and the error estimate for the Gauss-Legendre quadrature (cf. Lemma 1) then provides the bound

$$| (I^{(k)} - Q^{(k)}) v(\mathbf{y}_k^*) | \leq g(\tau_k) \exp \left( -h(\tau_k) \left( 2N_k^{(l)} - 1 \right) \right) \sup_{z \in \Sigma(\Gamma_k, \tau_k)} |v(z, \mathbf{y}_k^*)|$$

with  $g(\tau) = \frac{4}{(2\tau + \sqrt{1+4\tau^2-1})}$ . Inserting the choice of  $N_k^{(l)}$  yields

$$| (I^{(k)} - Q^{(k)}) v(\mathbf{y}_k^*) | \leq g(\tau_k) 2^{-l} \sup_{z \in \Sigma(\Gamma_k, \tau_k)} |v(z, \mathbf{y}_k^*)|.$$

Therefore,

$$|(I - Q)v| \leq 2^{-l} \sum_{k=1}^m g(\tau_k) \sup_{z \in \Sigma(\Gamma_k, \tau_k)} \sup_{\mathbf{y}_k^* \in \Gamma_k^*} |v(z, \mathbf{y}_k^*)|.$$

□

Since  $g(\tau) \lesssim \tau^{-1}$ , the error estimate in Lemma 7 is independent of  $m$  if the sequence  $\{\tau_k^{-1}\}_k$  is summable. Furthermore, for the assumptions to be satisfied,  $G$  must be analytically extendable. Therefore, we assume the following:

**Assumption 2.** Let  $G \in C^0(\Gamma)$  admit the partial derivatives  $\partial_{y_k}^j G$  for all  $k \in \mathbb{N}$  and let the following bounds hold for all  $k \in \mathbb{N}$

$$\|\partial_{y_k}^j G(\mathbf{y})\|_{C^0(\Gamma)} \lesssim j! \gamma_k^j.$$

We further assume either  $\gamma_k = k^{-\alpha}$  for some  $(1 - \delta)\alpha > 1$  with  $\delta \in (0, 1)$  fixed or  $\gamma_k = 2^{-\beta k}$  for some  $\beta > 0$ .

By Lemma 2, the Assumption 2 implies that for each  $k \in \mathbb{N}$  the function  $G$  can be extended into  $\Sigma(\Gamma_k, \tau_k) = \{z \in \mathbb{C} : \text{dist}(z, \Gamma_k) \leq \tau_k\}$  with  $\tau_k \in (0, \gamma_k^{-1})$ . So we can extend  $G$  into  $\Sigma(\Gamma_k, \tau_k)$  with  $\tau_k = \gamma_k^{-1+\delta}$  for some fixed  $\delta \in (0, 1)$ . We note that the sequence  $\{\tau_k\}$  grows to infinity for  $k \rightarrow \infty$ . The introduction of  $\delta > 0$  ensures that  $\gamma_k \tau_k = \gamma_k^\delta \rightarrow 0$  as  $k \rightarrow \infty$ , which will be needed later for the combination with different truncation levels. The parameter  $\delta$ , however, increases the cost of the product quadrature. The assumptions on the form of  $\gamma_k$  provides the summability of  $\{\tau_k^{-1}\}_k$ .

Having an estimate for the error, we also want a dimension-independent bound regarding the cost complexity of the product quadrature.

**Lemma 8.** Let the  $N_k^{(l)}$  be chosen according to

$$N_k^{(l)} = \left\lceil \frac{\ln(2)}{2h(\tau_k)}l + \frac{1}{2} \right\rceil.$$

Then the total number of quadrature points used for  $\mathbf{Q}_l$  can be bounded

(i) in the case of  $\gamma_k \leq 2^{-\beta k}$  by

$$\text{cost}(\mathbf{Q}_l) \lesssim 2^{\frac{1}{3(1-\delta)\beta}l \log_2(\frac{l-2}{(1-\delta)\beta} + 1) + \frac{\kappa}{(1-\delta)\beta}l}$$

with  $\kappa = \log_2(\frac{3}{2}e^{1/3})$ ,

(ii) and for  $\gamma_k \leq k^{-\alpha}$  by

$$\text{cost}(\mathbf{Q}_l) \lesssim 2^{\kappa 2^{l/(1-\delta)\alpha} + \vartheta l}$$

with  $\kappa = \log_2(\frac{3}{2}e^{1/3})$  and  $\vartheta = \frac{1}{3\ln(2)} \left( \frac{1}{2} + \frac{1}{2+(1-\delta)\alpha} \right)$ .

*Proof.* The cost complexity of  $\mathbf{Q}_l$  is simply the product of the number of quadrature points in each direction.

$$\text{cost}(\mathbf{Q}_l) = \prod_{k=1}^m N_k^{(l)}.$$

As  $m$  is not fixed and we want a bound independent of  $m$ , we can view the function as a function on  $\Gamma = \times_{k=1}^{\infty} \Gamma$  and bound the cost by

$$\text{cost}(\mathbf{Q}_l) \leq \prod_{k=1}^{\infty} N_k^{(l)}.$$

For a fixed  $l$ , the number of  $N_k^{(l)}$  decreases with increasing  $k$  and eventually  $N_k^{(l)} = 1$ . Hence, the infinite product is well-defined and the cost complexity is given only in terms of  $l$ .

In the following, we use the behavior of  $\tau_k$  to identify the number of terms to be considered in the product. For clarity, we assume the constant in  $\gamma_k \lesssim k^{-\alpha}$  and  $\gamma_k \lesssim 2^{-\beta k}$  to equal 1.

(i) First, the case  $\gamma_k \lesssim 2^{-\beta k}$  is considered. For fixed  $l$ ,  $N_k^{(l)} > 1$  if and only if  $\frac{\ln(2)}{2h(\tau_k)}l > \frac{1}{2}$ . This holds true when

$$l \ln(2) > h(\tau_k) = \ln \left( 2\tau_k + \sqrt{1 + 4\tau_k^2} \right) > \ln(4\tau_k).$$

With  $\tau_k \geq 2^{(1-\delta)\beta k}$  it follows

$$l > 2 + (1 - \delta)\beta k$$

which is equivalent to

$$k < \frac{l-2}{(1-\delta)\beta}.$$

We define  $M_l = \lceil \frac{l-2}{(1-\delta)\beta} \rceil$  such that  $N_k^{(l)} = 1$  for  $k > M_l$ . The cost is then estimated by

$$\begin{aligned} \text{cost}(\mathbf{Q}_l) &= \prod_{k=1}^{\infty} N_k^{(l)} = \prod_{k=1}^{M_l} \left[ \frac{\ln(2)l}{2(\ln(2\tau_k + \sqrt{1+4\tau_k^2}))} + \frac{1}{2} \right] \leq \prod_{k=1}^{M_l} \frac{\ln(2)l}{2\ln(4\tau_k)} + \frac{3}{2} \\ &\leq \prod_{k=1}^{M_l} \frac{l}{2((1-\delta)\beta k + 2)} + \frac{3}{2} = \prod_{k=1}^{M_l} \frac{3}{2} \left( \frac{l}{3((1-\delta)\beta k + 2)} + 1 \right). \end{aligned}$$

In order to estimate the product, we can consider the logarithm of the cost

$$\begin{aligned} \ln(\text{cost}(\mathbf{Q}_l)) &= \sum_{k=1}^{M_l} \left( \ln\left(\frac{3}{2}\right) + \ln\left(\frac{l}{3((1-\delta)\beta k + 2)} + 1\right) \right) \\ &\leq M_l \ln\left(\frac{3}{2}\right) + \sum_{k=1}^{M_l} \frac{l}{3((1-\delta)\beta k + 2)}, \end{aligned}$$

where we used that  $\ln(x+1) < x$  for  $x > 0$ .

The sum can be bounded above by

$$\sum_{k=1}^{M_l} \frac{l}{3((1-\delta)\beta k + 2)} \leq \frac{l}{3(1-\delta)\beta} \sum_{k=1}^{M_l} \frac{1}{k} \leq \frac{l}{3(1-\delta)\beta} (1 + \ln(M_l)).$$

Hence, for the cost it holds, with the estimate  $M_l \leq \frac{l-2}{(1-\delta)\beta} + 1$ ,

$$\begin{aligned} \text{cost}(\mathbf{Q}_l) &\leq \exp\left( M_l \ln\left(\frac{3}{2}\right) + \frac{l}{3(1-\delta)\beta} (1 + \ln(M_l)) \right) \\ &\leq 2^{\frac{l}{(1-\delta)\beta} \log_2\left(\frac{3}{2}\right) + \frac{\log_2(e)}{3(1-\delta)\beta} l + \frac{l}{3(1-\delta)\beta} \log_2\left(\frac{l-2}{(1-\delta)\beta} + 1\right)} \\ &= 2^{\frac{1}{3(1-\delta)\beta} l \log_2\left(\frac{l-2}{(1-\delta)\beta} + 1\right) + \frac{\kappa}{(1-\delta)\beta} l} \end{aligned}$$

with  $\kappa = \log_2\left(\frac{3}{2}e^{1/3}\right)$ .

(ii) Similarly, for the case  $\gamma_k \leq k^{-\alpha}$ , we obtain  $N_k^{(l)} > 1$  only if

$$k < 2^{\frac{l-2}{(1-\delta)\alpha}}.$$

Thus, we set  $M_l = \lceil 2^{\frac{l-2}{(1-\delta)\alpha}} \rceil$  and apply similar arguments as before.

$$\prod_{k=1}^{\infty} N_k \leq \prod_{k=1}^{M_l} \frac{\ln(2)}{2\ln(4k^{(1-\delta)\alpha})} l + \frac{3}{2} = \prod_{k=1}^{M_l} \frac{3}{2} \left( \frac{\ln(2)l}{3(\ln(4) + (1-\delta)\alpha \ln(k))} + 1 \right)$$

and for the logarithm of the cost

$$\ln(\text{cost}(\mathbf{Q}_l)) \leq M_l \ln\left(\frac{3}{2}\right) + \sum_{k=1}^{M_l} \frac{\ln(2)l}{3(\ln(4) + (1-\delta)\alpha \ln(k))}.$$

The sum can be bounded by

$$\begin{aligned} & \sum_{k=1}^{M_l} \frac{\ln(2)l}{3(\ln(4) + (1-\delta)\alpha \ln(k))} \\ & \leq \frac{\ln(2)l}{3} \left( \frac{1}{\ln(4)} + \frac{1}{\ln(4) + (1-\delta)\alpha \ln(2)} + \sum_{k=3}^{M_l} \frac{1}{\ln(4) + (1-\delta)\alpha \ln(k)} \right) \\ & \leq \frac{l}{3} \left( \frac{1}{2} + \frac{1}{2 + (1-\delta)\alpha} \right) + \frac{\ln(2)l}{3(1-\delta)\alpha} \int_2^{M_l} \frac{1}{\ln(x)} dx \\ & \leq \frac{l}{3} \left( \frac{1}{2} + \frac{1}{2 + (1-\delta)\alpha} + \frac{\ln(2)}{(1-\delta)\alpha} \text{Li}(M_l) \right) \end{aligned}$$

where  $\text{Li}(x)$  is the logarithmic integral function. It is known that  $\text{Li}(x)$  behaves asymptotically like  $\mathcal{O}\left(\frac{x}{\ln(x)}\right)$  with a constant approaching 1 as  $x \rightarrow \infty$ . Hence, together with the bound  $M_l \leq 2^{l/(1-\delta)\alpha}$  at least for  $l$  large enough, we get

$$\text{Li}(M_l) \leq \frac{(1-\delta)\alpha 2^{l/\alpha}}{\ln(2)l}.$$

Altogether,

$$\begin{aligned} \text{cost}(\mathbf{Q}_l) & \leq \exp\left(2^{l/(1-\delta)\alpha} \ln\left(\frac{3}{2}\right) + \frac{l}{3} \left(\frac{1}{2} + \frac{1}{2 + (1-\delta)\alpha}\right) + \frac{1}{3} 2^{l/(1-\delta)\alpha}\right) \\ & \lesssim 2^{\kappa 2^{l/(1-\delta)\alpha} + \vartheta l} \end{aligned}$$

$$\kappa = \log_2\left(\frac{3}{2}e^{1/3}\right) \text{ and } \vartheta = \frac{1}{3\ln(2)} \left(\frac{1}{2} + \frac{1}{2+(1-\delta)\alpha}\right).$$

□

In both cases, the cost complexity of the product quadrature grows more than exponentially in  $l$ . The actual cost might be lower when  $m < M_l$ . But this leads to a bound which depends on  $m$ . Even an exponential decay of the influence on  $G$  of the variables  $y_k$  is not enough for the cost being of the form  $\mathcal{O}(2^{sl})$ ,  $s > 0$ . For that to hold,  $\{\gamma_k\}$  would need to converge faster than exponentially.

From Lemma 8 it is clear that for larger  $\alpha$  and  $\beta$  the product quadrature is less expensive. This can be expected as the higher decay in importance of the variables is transferred to the choice of number of points  $N_k^{(l)}$  and this requires less dimensions with  $N_k^{(l)} > 1$ .

Compared to the Monte Carlo quadrature, the anisotropic Gaussian product quadrature does not provide the improvement which was hoped to be achieved.

$\alpha$	2	3	4	5	6
$l$	5	12	18	26	33

$\beta$	0.5	0.75	1	1.25	1.5
$l$	1	2	7	25	84

Table 2: Maximal level  $l$  for which the cost in order to achieve an error of  $\mathcal{O}(2^{-l})$  is lower for the product quadrature than for the Monte Carlo method. The maximal levels for different decay rates are shown which are not slightly lowered by  $\delta$  (i.e.,  $\delta = 0$ ).

For the same order of error, say  $\mathcal{O}(2^{-l})$ , the Monte Carlo method needs  $\mathcal{O}(2^{2l})$  sample points, which is less than the product quadrature, at least asymptotically.

However, for small and moderate levels  $l$  the required cost might be smaller for the product quadrature. This regime, though, depends on the constants  $\alpha$  and  $\beta$  and the actual constant in the estimates  $\gamma_k \lesssim k^{-\alpha}$  or  $\gamma_k \lesssim 2^{-\beta k}$ , respectively. For an illustration, the maximal level  $l$  where the product quadrature is more profitable than the Monte Carlo method are tabulated in table 2 for different  $\alpha$  and  $\beta$  under the assumption that all constants equal 1.

### Sparse Grid

Next, we focus on how this product quadrature can be used in a sparse grid method together with the truncation level of  $G$ . We need, hereby, to satisfy the conditions (41) and (38). The latter is provided by Lemma 7 under the Assumption 2 on  $G$  with  $s_1 = 1$  and the norm  $\|\cdot\|_{\mathcal{Y}}$  defined by

$$\|v\|_{\mathcal{Y}} = \sup_k \|v\|_{C^0(\Sigma(\Gamma_k, \tau_k); C^0(\Gamma_k^*))} = \sup_k \sup_{y_k \in \Sigma(\Gamma_k, \tau_k)} \sup_{\mathbf{y}_k^* \in \Gamma_k^*} |v(y_k, \mathbf{y}_k^*)|.$$

Hence, the only point left is to verify that the truncation levels  $m_l$  are chosen such that the condition (41) is satisfied. For that, we need a stronger assumption on  $G$  than Assumption 2. The bound on the first derivatives must also hold in the region of analytic extension. This is similar to requiring a mixed condition for the sparse grid.

**Assumption 3.** *Let the function  $G$  and the sequence  $\{\gamma_k\}_k$  be such that Assumption 2 is satisfied. In addition, assume that the bound for the first derivative hold also on the extension into the complex plane, i.e., for any  $k \in \mathbb{N}$*

$$\sup_{y_k \in \Sigma(\Gamma_k, \tau_k)} \sup_{\mathbf{y}_k^* \in \Gamma_k^*} |\partial_{y_i} G(y_k, \mathbf{y}_k^*)| \lesssim \gamma_i$$

with a constant independent of  $k$ .

**Lemma 9.** *Under Assumption 3 condition (41) is satisfied, i.e.,*

$$\|G(\mathbf{y}) - G_{m_l}(\mathbf{y})\|_{\mathcal{Y}} \lesssim 2^{-s_2 l}$$

(i) with  $s_2 = \delta\alpha$  if  $m_l = 2^l$  in the case of  $\gamma_k \lesssim k^{-\alpha}$ .

(ii) with  $s_2 = \delta\beta$  if  $m_l = l$  in the case  $\gamma_k \lesssim 2^{-\beta k}$ .

*Proof.* For any  $k$  consider

$$\sup_{y_k \in \Sigma(\Gamma_k, \tau_k)} \sup_{\mathbf{y}_k^* \in \Gamma_k^*} |G(y_k, \mathbf{y}_k^*) - G_{m_l}(y_k, \mathbf{y}_k^*)|.$$

Since  $|G(y_k, \mathbf{y}_k^*) - G_{m_l}(y_k, \mathbf{y}_k^*)| \leq \sum_{i=m_l+1}^{\infty} \sup_{t_k \in \Sigma(\Gamma_k, \tau_k)} \sup_{\mathbf{t}_k^* \in \Gamma_k^*} |\partial_i G(t_k, \mathbf{t}_k^*)| |y_i|$ , the Assumption 3 implies for  $k \geq m_l + 1$

$$\begin{aligned} & \sup_{y_k \in \Sigma(\Gamma_k, \tau_k)} \sup_{\mathbf{y}_k^* \in \Gamma_k^*} |G(y_k, \mathbf{y}_k^*) - G_{m_l}(y_k, \mathbf{y}_k^*)| \\ & \leq \frac{1}{2} \sum_{i=m_l+1, i \neq k}^{\infty} \gamma_i + \gamma_k \left( \frac{1}{2} + \gamma_k^{-(1-\delta)} \right) \\ & \lesssim \sum_{i=m_l+1}^{\infty} \gamma_i + \gamma_k^{\delta}. \end{aligned} \tag{46}$$

Since  $\gamma_k$  is decreasing, the largest value is obtained for  $k = m_l + 1$ .

The sum can be estimated in the same manner as in the proof of Lemma 6 by  $\sum_{i=m_l+1}^{\infty} \gamma_i \lesssim 2^{-(1-\alpha)l}$  or  $\sum_{i=m_l+1}^{\infty} \gamma_i \lesssim 2^{-\beta l}$ . Hence, in (46), the second term dominates the bound

$$\|G - G_{m_l}\|_{\mathcal{Y}} \lesssim \begin{cases} 2^{-\delta\alpha l} \\ 2^{-\delta\beta l}. \end{cases}$$

□

This provides the necessary convergence rate 41 of the truncation for the product quadrature such that Theorem 7 holds. However, as the computational cost grows fast for  $\mathbf{Q}_l$ , we cannot assess the computational cost of the sparse grid in the same way as in Theorem 8. Eventually, the cost of the product quadrature will dominate the cost of the sparse grid. But since the quadrature rule is applied to a finite-dimensional integrand, the bound provided in Theorem 8 is only an upper bound. In the numerical examples, we find that the sparse grid with the product quadrature yields a better convergence with respect to the cost than the Monte Carlo quadrature, at least for a moderate overall sparse grid level  $L$ .

## Connection to PDE

Before looking at numerical examples, we want to connect the ideas above to the context of computing the first and second moment for a PDE solution where the diffusion is given by a Karhunen-Loeve expansion of a random field. As mentioned before, this problem involves an infinite-dimensional parameter space.

In this case, the function  $G$  corresponds to the map  $\mathbf{y} \mapsto u(\mathbf{y})$  or  $\mathbf{y} \mapsto u(\mathbf{y})^2$ , where  $u(\mathbf{y})$  is the solution to the PDE

$$-\operatorname{div}(a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y})) = f(\mathbf{x}).$$

In contrast to the above description of the integration problem, the function takes values in a function space, like  $H_0^1(D)$  instead of  $\mathbb{R}$ . Hence, the integral is a Bochner integral and the assumptions must hold with respect to the Bochner space.

In this case, the anisotropy of the integrand with respect to the stochastic variables is provided by the anisotropic dependence of the random field on the variables. The sequence  $\{\gamma_k\}_k$  is given by  $\gamma_k = \tilde{\gamma}_k/a_{\min}$  where  $\tilde{\gamma}_k$  expresses the decay of the coefficients in the Karhunen-Loeve expansion of the random field  $a(\mathbf{x}, \mathbf{y})$ , i.e.,

$$\tilde{\gamma}_k = \sqrt{\lambda_k} \|\psi_k(\mathbf{x})\|_{L^\infty(D)},$$

where  $\lambda_k$  are the eigenvalues and  $\psi_k(\mathbf{x})$  the eigenfunctions. As discussed in Section 3.3, the sequence satisfies certain decay rates depending on the regularity of the covariance function, which are either algebraic for example for the Matérn covariance or exponential for a Gaussian covariance function. This explains why we considered these choices for  $\gamma_k$ .

In [57], it was proven that the anisotropy of the field is transferred to the solution of the PDE. A similar result like Lemma 5 provides the analytic extendibility of the PDE solution.

**Lemma 10.** *For any sequence of nonnegative integers  $\boldsymbol{\nu} = (\nu_j)_{j \geq 1}$  with only finitely many  $\nu_j$  nonzero, the function  $\mathbf{y} \mapsto u(\mathbf{y})$  admits a partial derivative  $\partial_{\mathbf{y}}^{\boldsymbol{\nu}}$  and for all  $\mathbf{y} \in \Gamma$*

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} u(\mathbf{y})\|_{H_0^1(D)} \lesssim |\boldsymbol{\nu}|! \left( \frac{\tilde{\gamma}}{a_{\min}} \right)^{\boldsymbol{\nu}}.$$

By this, the Assumption 2 is guaranteed. Furthermore, it has been shown that the dependence of the solution on the parameters is analytically smooth [58], such that the Assumption 3 holds by similar arguments as for Lemma 10.

Up to this point, the spatial discretization has not been considered. However, the PDE can only be solved numerically, so that the assumptions need to be verified also for the spatial approximate solution. Furthermore, the spatial discretization error needs to be balanced with the truncation and stochastic approximation error. A further step for the computation of moments would therefore be to include the spatial discretization in the sparse grid. We will not address this here. Instead, an a-posteriori approach will be discussed in the second part. For a given problem with countably many stochastic parameters, the method that will be described determines adaptively how to balance the different directions (see Section 6.2).

### 5.2.3 Numerical Examples

This section is dedicated to numerical results which illustrate the above conclusions. As an example for  $G$ , we look at a function which is of the form

$$G(\mathbf{y}) = F \left( c_0 + \sum_{k=1}^{\infty} \gamma_k y_k \right),$$

where  $F$  is a holomorphic function and  $\gamma_k$  fulfills the above assumptions. We choose the functions

$$G_1(\mathbf{y}) = \frac{1}{2 + \sum_{k=1}^{\infty} k^{-2} y_k}$$

and

$$G_2(\mathbf{y}) = \frac{1}{6 + \sum_{k=1}^{\infty} 2^{-k/2} y_k}$$

for our examples.

### Numerical Examples Monte Carlo

First, we illustrate the motivation behind using a sparse grid between the truncation level and quadrature method. For that, we look at the benefit-cost ratio for the combination of truncating to a finite-dimensional space and applying the Monte Carlo quadrature. The result for  $G_1$  is shown in Figure 8. The decrease of the benefit-cost ratio and the shape of a scalene triangle are clearly visible. The slope of the contour lines is  $3/4$  as expected by the convergence rates of each method separately. Therefore, including the hierarchical increments in the triangle set with  $\sigma = \sqrt{3/4}$  reduces the cost compared to a full grid solution in an optimal way.

Convergence rates for this choice of  $\sigma$ , based on an equilibrated benefit-cost ratio, are shown in Figure 9, whereby the errors are measured with respect to a reference solution obtained by the full combination of quadrature and truncation level. As the rates only hold in the root mean square sense, the RMSE calculated over five realizations and three realizations are shown. As a reference, we also display the predicted rates given by Theorem 7 and Corollary 2. We observe that they are achieved, but the Monte Carlo method limits the convergence to the rate  $1/2$ .

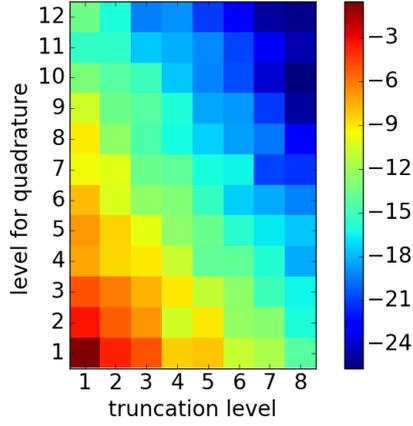


Figure 8: Tableau of benefit-ratio with  $N_y = 2^l$  quadrature points and truncation levels  $m_l = 2^l$  in a logarithmic plot.

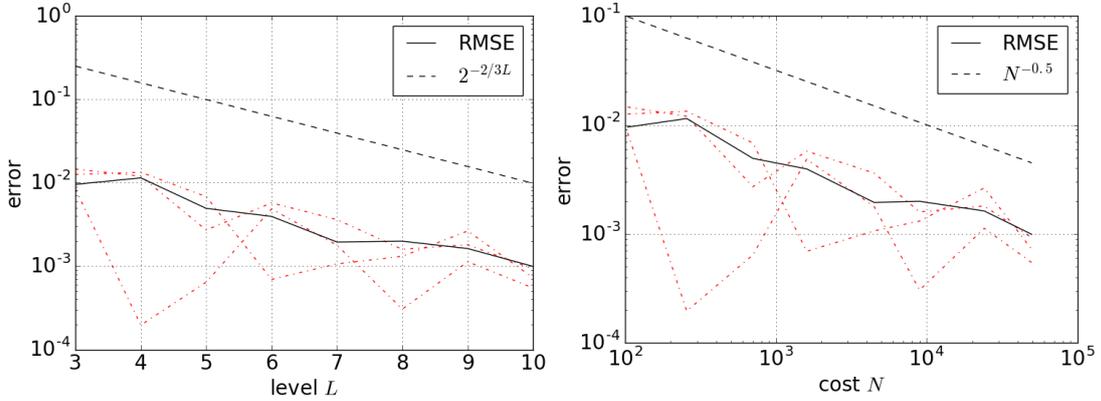


Figure 9: Convergence behavior of the sparse grid with a Monte Carlo and truncation with respect to the sparse grid level (left) and computational cost (right). The parameter  $\sigma = \sqrt{3/4}$  is chosen to equilibrate the benefit-cost ratio and the error is measured in RMSE sense with five realizations. The dotted lines (red) show three realizations.

For the case where  $\gamma_k$  decays exponentially, we consider function  $G_2$ . The results are illustrated in Figure 10 for different choices of  $\sigma$ . They again show the expected convergence rates. Furthermore, for  $\sigma^2 < 1$  the rate of convergence is less than  $1/2$  as predicted. Nevertheless, a larger choice for  $\sigma$  does not improve the convergence rate to more than  $1/2$  in terms of computational cost, as the limiting factor then becomes the Monte Carlo method.

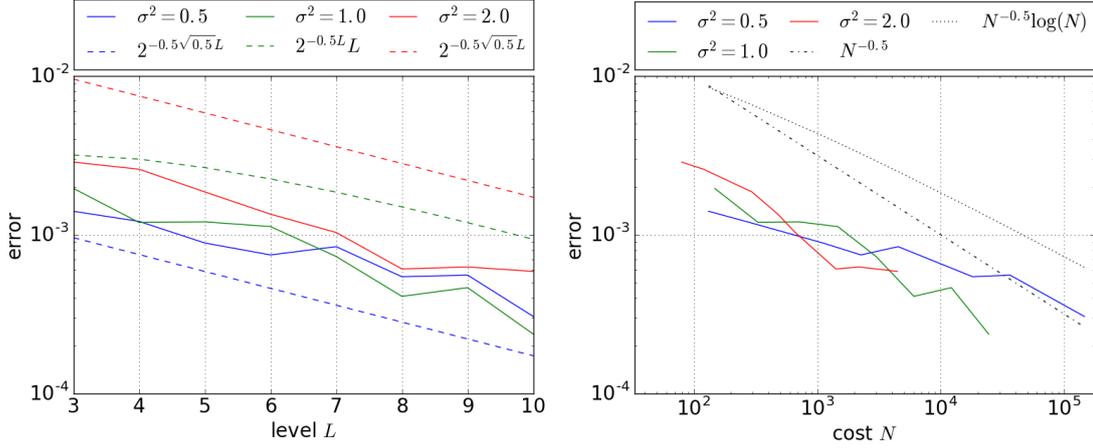


Figure 10: Convergence of the RMSE with respect to sparse grid level and cost for  $G_2$  with different choices for  $\sigma$ . The expected rates are plotted with dotted lines.

## Numerical Examples Product Quadrature

Next, we look at the product quadrature. First, we examine the convergence behavior of the product quadrature on its own by means of function  $G_1$ . Figure 11 shows the behavior of error and cost for different choices of  $\delta$ . The rate for the error  $\mathcal{O}(2^{-l})$  as given by Theorem 7 is achieved with only minor differences in  $\delta$ . An even better than expected decrease is observed, as, due to rounding up the number of quadrature points,  $N_k^{(l)}$  is slightly larger than required. For the computational cost, however, we observe that the estimates in Lemma 8 are too strong and the increase of the actual computational cost is slower. The bound for rounding the numbers  $N_k^{(l)}$  and further estimates might be lower in practice. Nevertheless, it can be observed that the real computational cost grows faster than geometrically for all choices of  $\delta$ . Since a larger choice of  $\delta$  increases the number of quadrature points, it is not surprising that this leads to a higher computational cost. We recall that  $\delta$  was only introduced for the sparse grid, so that we could set  $\delta = 0$  when only considering the product quadrature, which leads to the smallest cost.

The last panel in Figure 11 illustrates the convergence with respect to the actual computational cost for  $\delta = 0$ . Since the number of quadrature points increases faster than exponentially with level  $l$ , the error decreases first faster and then only slowly. The rate clearly surpasses the one of a Monte Carlo quadrature for the first levels, but then the product quadrature performs worse.

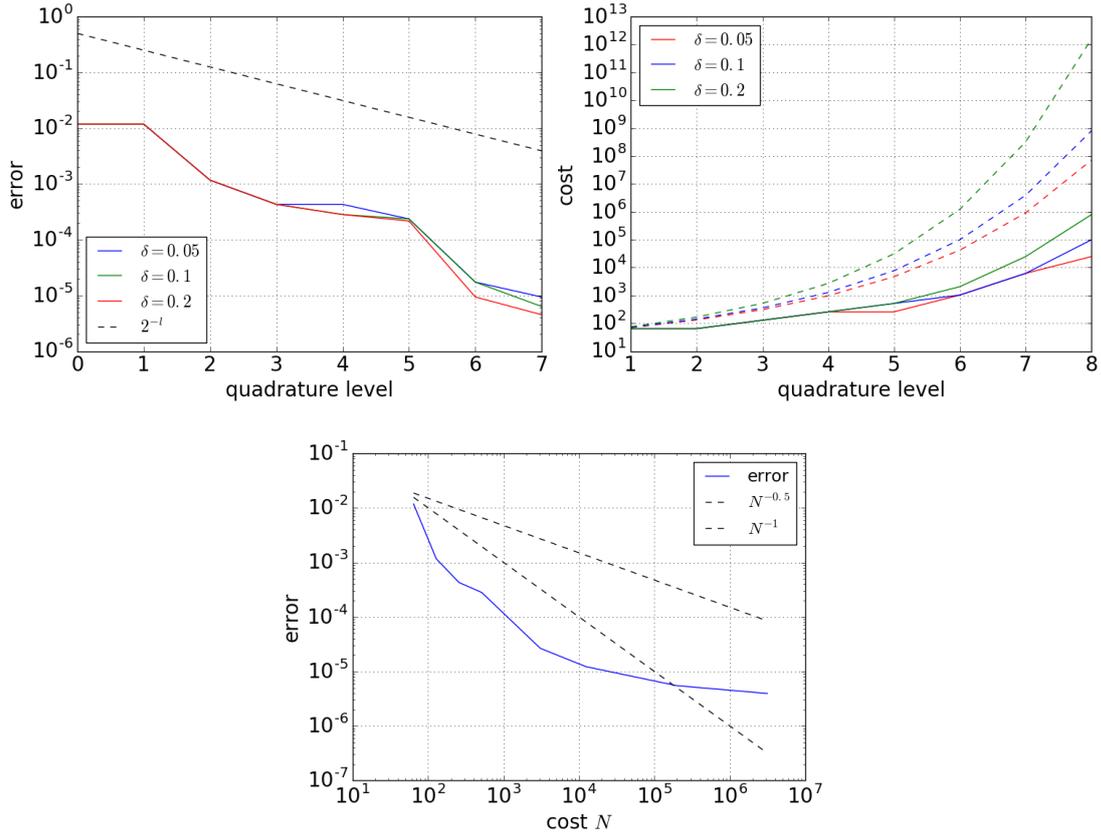


Figure 11: Convergence behavior for the product quadrature. Error measured with respect to a reference solution. Top left: error rate for different choices of  $\delta$  with respect to the discretization level, top right: actual cost (solid) and predicted upper bounds (dashed) for different  $\delta$ , bottom: error rate with respect to computational cost for  $\delta = 0$ .

Finally, we look at the combination technique using the product quadrature. As the results are similar for the case of faster convergent coefficients, we only present the case for function  $G_1$ . We fix the parameter  $\delta = 0.1$  and compute the errors with respect to a full combination reference solution  $\mathbf{Q}_{11}G_{1,m_8}$ . The left panel in Figure 12 shows the error with respect to the discretization level for the choice  $\sigma^2 = 5$ , which corresponds to an equilibrated approximation power. We see that the error decreases even faster than the expected rate from Theorem 7 for similar reasons as above.

The right panel illustrates the error in terms of computational cost for different  $\sigma$ . Since the error of the product quadrature does not decrease with a constant rate, the best choice for  $\sigma$  is not obvious. For moderate  $L$ , a small  $\sigma$  and, therefore, including higher levels for the quadrature method seems to work better. For the choice  $\sigma^2 = 0.5$ , a rate of roughly 1 is achieved, which is better than the Monte Carlo rate, whereas larger  $\sigma$  lead to smaller rates. Since for the anisotropic sparse grid the product quadratures with high levels are applied to functions where the variable space has been truncated to a small number, the computational cost for these quadratures is not as large as when applied to  $G_1(\mathbf{y})$ . The rate of

the cost increase in the quadrature is, therefore, rather based on the rate for lower discretization levels. This explains why  $\sigma < 1$  leads to better results. By this choice higher truncation levels than quadrature levels are considered, as the quadrature is less expensive for small levels.

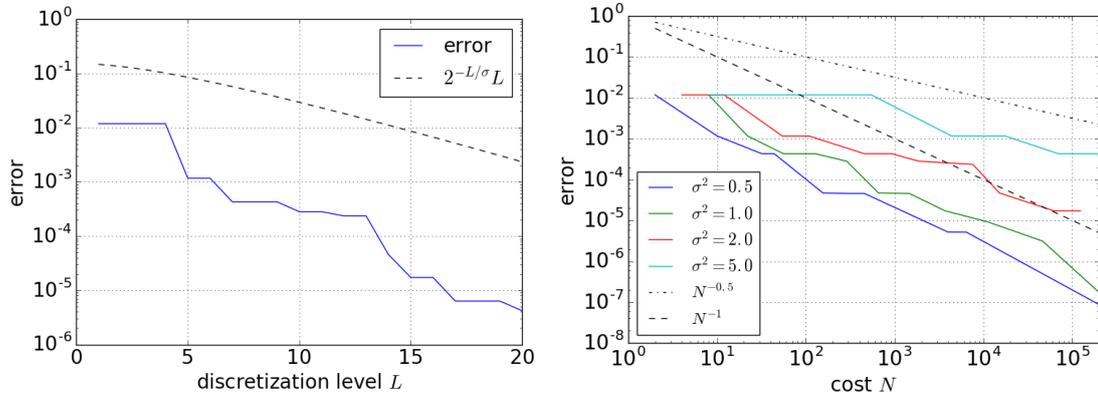


Figure 12: Anisotropic sparse grid with product quadrature. Left: convergence with respect to discretization level for  $\sigma^2 = 1/(2\delta)$  and  $\delta = 0.1$ . Right: convergence with respect to computational cost for different  $\sigma$  and  $\delta = 0.1$ .

## 6 Dimension-adaptive Combination Technique

In this part of the thesis, we will follow a different approach than before and look at an adaptive algorithm that constructs the sparse grid step by step. We will discuss here the dimension-adaptive combination technique which is an a-posteriori approach to the problem. Instead of applying a fixed structure of a sparse grid to the problem, the dimension-adaptive combination technique aims at identifying the important dimensions and refines adaptively the associated solvers, whereby this identification process is controlled by suitable indicators.

Different to the a-priori case, the dimension-adaptive combination technique does not require information about the solution beforehand. While in the previous sections the optimal way to balance the different numerical methods was determined based on the size of the contributions  $\Delta_l u$  which could be estimated by the regularity of the solution, the adaptive algorithm constructs heuristically the sparse grid. Hence, the dimension-adaptive combination technique can also be applied to problems where it is not clear how to combine the numerical approximations, for example, in the case of unsmooth solutions or insufficient knowledge about the regularity.

The dimension-adaptive combination technique has been applied in the literature in order to construct adaptively sparse grids in various contexts. The basic algorithm was developed by Gerstner and Griebel for the tensor product of quadrature rules [7]. This algorithm was adapted later to other problems including the approximation of PDEs (see, e.g., [59]).

Here, we adapt the algorithm in order to compute the first and second moment of the solution to the PDE (3). Similar algorithms have been explored for this type of problem of stochastic parameterized elliptic PDEs in [60] and [61]. These works, however, only consider a dimension-adaptive combination technique in the stochastic parameter space. In contrast to that, we design a dimension-adaptive combination technique such that a sparse grid that balances the quadrature rules and the spatial discretization is constructed. In addition, we extend the algorithm to the case where the diffusion coefficient in the PDE is given as a random field in the Karhunen-Loeve expansion and hence depends on infinitely many parameters. In this case, also the truncation level is adjusted during the algorithm. An adaptive algorithm where the effective stochastic dimension varies has been addressed in [61]. However, the truncation level was not balanced with the finite element method and quadrature rules. Here, we take these aspects into account by considering the truncation as an additional direction in the sparse grid formulation.

In the following, we first describe how the dimension-adaptive combination technique can be adapted to compute the quantities of interest in the case where the dimension of the parameter space is finite or fixed beforehand by a truncation level that cannot be changed. Later, in the second part, we focus on the case of an infinite-dimensional parameter space and construct an algorithm that adaptively determines the needed number of dimensions of the parameter space.

## 6.1 Finite-dimensional Parameter Space

In this section, we assume that the parameter space is  $m$ -dimensional and that the method for the spatial discretization is characterized by a single level parameter. For solving the PDE, this means we use a finite element method on a uniform mesh with the same mesh width in all spatial dimensions. Hence, we have  $m+1$  directions that have to be considered by the combination technique, one spatial and  $m$  stochastic dimensions. Given an array of levels  $\mathbf{l} \in \mathbb{N}^{m+1}$ , the first entry corresponds to the spatial level such that the solver uses a mesh with  $N_{l_1}^x = \mathcal{O}(2^{l_1+1})$  points in each direction. The remaining entries specify the stochastic levels with  $N_{l_{i+1}}^{y_i} = 2^{l_{i+1}}$  quadrature points in each dimension  $1 \leq i \leq m$ .

The dimension-adaptive algorithm constructs heuristically the set of increment spaces that are used for the sparse grid. Hence, the sparse grid formula is based on a generalized index set  $\mathcal{I}$ . Different to the a-priori case where the index set is known beforehand, the main goal of the adaptive algorithm is to find a good index set for the problem under consideration.

Given an index set  $\mathcal{I}$ , the sparse grid solution is obtained by including the increments for the indices contained in  $\mathcal{I}$

$$\mathbf{P}_{\mathcal{I}}u = \sum_{\mathbf{l} \in \mathcal{I}} \Delta_{\mathbf{l}}u \quad (47)$$

with  $\Delta_{\mathbf{l}}u = \Delta_{l_1}^{(1)} \otimes \dots \otimes \Delta_{l_d}^{(d)}u$  and the difference operators

$$\Delta_{l_i}^{(i)}u = \begin{cases} P_{l_i}^{(i)}u - P_{l_i-1}^{(i)}u & \text{if } l_i \geq 1, \\ P_{l_0}^{(i)}u & \text{if } l_i = 0, \end{cases}$$

as before. However, we want to use the full grid solutions to construct the approximate solution and therefore consider the representation by the combination technique formula which we stated in Section 4.2. For a generalized index set, we obtain the solution by

$$\mathbf{P}_{\mathcal{I}}^{\text{ct}}u = \sum_{\mathbf{l} \in \mathcal{I}} c_{\mathbf{l}} \mathbf{P}_{\mathbf{l}}u \quad (48)$$

with the coefficients

$$c_{\mathbf{l}} = \sum_{\mathbf{z}=\mathbf{0}}^{\mathbf{1}} (-1)^{|\mathbf{z}|_1} \chi^{\mathcal{I}}(\mathbf{l} + \mathbf{z}),$$

where  $\mathbf{1} = (1, \dots, 1)$  and the characteristic function  $\chi$  is given by

$$\chi^{\mathcal{I}}(\mathbf{l}) = \begin{cases} 1 & \text{if } \mathbf{l} \in \mathcal{I} \\ 0 & \text{otherwise.} \end{cases}$$

In order to have an equivalence of the combination technique (48) and the sparse grid formulation based on the increments (47), we need to impose a condition on

the index set  $\mathcal{I}$ . As the equivalence is based on the telescoping sum effect, we require that the index set  $\mathcal{I}$  is *downward closed*, meaning it satisfies the following admissibility condition:

For any  $\mathbf{l} \in \mathcal{I}$  it must hold that

$$\mathbf{l} - \mathbf{e}_j \in \mathcal{I} \quad \text{for } 1 \leq j \leq d \text{ with } l_j \geq 1 \quad (49)$$

with  $\mathbf{e}_j$  being the  $j$ -th unit vector.

After having discussed how to obtain the solution provided a general index set, we address in the following how a suitable index set can be found. The procedure is described in Algorithm 1.

---

**Algorithm 1** Dimension-adaptive Combination Technique

---

**Input:** Tolerance  $\varepsilon > 0$

**Output:** Index set  $\mathcal{I}$  and solution  $\mathbf{P}_{\mathcal{I}}^{\text{ct}}u$

```

1:  $\mathbf{l} = (0, \dots, 0)$ 
2:  $\mathcal{A} = \{\mathbf{l}\}, \mathcal{O} = \emptyset$ 
3: compute local profit indicator  $\eta_{\mathbf{l}}$ 
4:  $\eta = \eta_{\mathbf{l}}$ 
5: while  $\eta > \varepsilon$  do
6:   Select the index  $\mathbf{j} \in \mathcal{A}$  with the largest profit  $\eta_{\mathbf{j}}$ 
7:    $\mathcal{A} = \mathcal{A} \setminus \{\mathbf{j}\}, \mathcal{O} = \mathcal{O} \cup \{\mathbf{j}\}$ 
8:    $\eta = \eta - \eta_{\mathbf{j}}$ 
9:   for  $k = 1, \dots, d$  do
10:     $\mathbf{l} = \mathbf{j} + \mathbf{e}_k$ 
11:    if  $\mathbf{l} - \mathbf{e}_i \in \mathcal{O} \quad \forall 1 \leq i \leq d$  then
12:       $\mathcal{A} = \mathcal{A} \cup \{\mathbf{l}\}$ 
13:      Compute full grid solution  $\mathbf{P}_{\mathbf{l}}u$ 
14:      Compute local profit indicator  $\eta_{\mathbf{l}}$ 
15:       $\eta = \eta + \eta_{\mathbf{l}}$ 
16:    end if
17:   end for
18: end while
19: Combine solutions  $\mathbf{P}_{\mathbf{l}}u$  for  $\mathbf{l} \in \mathcal{I} = \mathcal{A} \cup \mathcal{O}$  according to (48) to obtain  $\mathbf{P}_{\mathcal{I}}^{\text{ct}}u$ 
20: return  $\mathcal{I} = \mathcal{A} \cup \mathcal{O}$  and  $\mathbf{P}_{\mathcal{I}}^{\text{ct}}u$ 

```

---

The adaptive algorithm constructs a problem dependent sequence of index sets  $\mathcal{I}^{(1)} \subset \mathcal{I}^{(2)} \subset \dots \subset \mathcal{I}^{(it)}$  which are at all times admissible in the sense of (49). The algorithm starts off with one index in  $\mathcal{I}$ , namely  $\mathbf{l} = \mathbf{0}$ , and adds successively indices. To ensure the admissibility, we introduce two index sets  $\mathcal{A}$  and  $\mathcal{O}$  with  $\mathcal{I} = \mathcal{A} \cup \mathcal{O}$  and  $\mathcal{A} \cap \mathcal{O} = \emptyset$  at all times. The set  $\mathcal{A}$  denotes the active indices

and we refer to the set  $\mathcal{O}$  as the set of old indices. While the indices in  $\mathcal{O}$  have already been chosen to be considered for the sparse grid, the index set  $\mathcal{A}$  contains the indices in the neighborhood of  $\mathcal{O}$  which do not destroy the admissibility when added to  $\mathcal{O}$ . Hence, they form the set of indices which can be added in order to enlarge the set  $\mathcal{O}$ .

In each iteration of the dimension-adaptive combination technique, the index with the highest contribution is selected among the indices in  $\mathcal{A}$ . This selection is based on a local profit indicator  $\eta_l$  which will be discussed later on in more detail. The selected index is then removed from  $\mathcal{A}$  and added to the set  $\mathcal{O}$ . The neighborhood for the index set  $\mathcal{O}$  has changed as a result of this process such that there might be new indices, that if added to  $\mathcal{O}$ , the resulting set is still admissible in the sense of (49). Hence, the active index set  $\mathcal{A}$  needs to be updated by adding these indices to  $\mathcal{A}$ .

In a final step, the full grid solutions are combined according to (48). Since during the algorithm the solutions for  $\mathbf{l} \in \mathcal{A}$  were computed for the profit indicator, also these indices are included into the computation of the combination technique solution. The solution is built therefore with the index set  $\mathcal{I} = \mathcal{A} \cup \mathcal{O}$ , instead of  $\mathcal{O}$ .

A crucial point of the adaptive method is to select in line 6 the next index which is added to  $\mathcal{O}$ , as this step controls which direction is refined next. We estimate how much an index contributes to the solution by associating to each index  $\mathbf{l}$  that might be added to  $\mathcal{I}$  a local error indicator. Adding  $\mathbf{l}$  in  $\mathcal{I}$  corresponds to including the increment  $\Delta_{\mathbf{l}}$  in the sparse grid formula (47). Hence, we use as a local error indicator

$$E_{\mathbf{l}} = \|\Delta_{\mathbf{l}}u\|_{\mathcal{X}} = \left\| \left( \Delta_{l_1}^x \otimes \Delta_{l_2}^{y_1} \otimes \dots \otimes \Delta_{l_{m+1}}^{y_m} \right) u \right\|_{\mathcal{X}} \quad (50)$$

where  $\mathcal{X}$  corresponds to the space in which the approximation of the quantities of interest is considered. For evaluating  $E_{\mathbf{l}}$ , when  $\mathbf{l}$  is added,  $2d$  full grid solutions are required. However, we only need to compute  $\mathbf{P}_{\mathbf{l}}u$  and can use the full grid solutions that have been computed before in the course of the algorithm as the index set needs to be admissible.

Taking into account the computational cost, the local error indicator  $E_{\mathbf{l}}$  might not be an optimal refinement indicator, as including other indices might lead to a similar error reduction while having much lower computational cost. A further refinement in one direction only makes sense if the required work is not too large. Similar to the a-priori construction of the optimal index set for the regular sparse grid, we therefore consider the benefit-cost ratio as an indicator for the refinement of the numerical methods.

We measure the cost of computing the full grid solution  $\mathbf{P}_{\mathbf{l}}u$  in terms of the number of grid points used, i.e.,

$$c_{\mathbf{l}} = (N_{l_1}^x)^n \cdot \prod_{i=1}^m N_{l_{i+1}}^{y_i}$$

with  $n$  being the dimension of the spatial domain and define as local profit indicator

$$\eta_{\mathbf{l}} = \frac{E_{\mathbf{l}}}{c_{\mathbf{l}}}.$$

Consequently, the algorithm searches among the indices in  $\mathcal{A}$  for the one with the highest benefit-cost ratio and includes the indices for which either the error indicator is large or the cost small.

The limit of the dimension-adaptive combination technique, however, also becomes apparent from this refinement strategy. While in the a-priori case the regularity assumptions provided that the contribution of  $\Delta_{\mathbf{l}}u$  to the solution decreases with each refinement, we cannot ensure this for general problems. Hence, there may be forward neighbors of  $\mathbf{l}$  with a high profit indicator, although the indicator for  $\mathbf{l}$  is small. In this case the algorithm does not refine this direction adequately. But only an a-priori analysis can provide the knowledge that is needed to adjust the algorithm in order to circumvent these issues. In general, the consideration of the computational cost in the indicator already counteracts the problem, as not only the local error determines the refinement.

Apart from the refinement strategy, the stopping criterion must be chosen. As a stopping criterion for the algorithm, different criteria can be used which will be looked at later in the numerical examples. One option is to stop when the given upper bound on the work is exhausted:

$$C = \sum_{\mathbf{l} \in \mathcal{I}} c_{\mathbf{l}}.$$

Another is based on the profits in the neighborhood set  $\mathcal{A}$ . This approach estimates how much profit is not covered yet. As the error contribution for indices in the complement of  $\mathcal{I}$  is not known, the only available information are the contributions of the indices in  $\mathcal{A}$ . Therefore, an estimate of the global profit is

$$\eta = \sum_{\mathbf{l} \in \mathcal{A}} \eta_{\mathbf{l}} \tag{51}$$

The algorithm stops when the global profit indicator is below a tolerance  $\varepsilon$ . This stopping criterion is used in Algorithm 1. It is also possible to replace the profit by the error contribution in the stopping criterion providing a global error estimator.

A third stopping criterion looks at the difference between the solutions of two iterations in order to estimate the error, i.e.,

$$\|\mathbf{P}_{\mathcal{I}^{(t+1)}}^{\text{ct}} u - \mathbf{P}_{\mathcal{I}^{(t)}}^{\text{ct}} u\|_{\mathcal{X}} \leq \sum_{\mathbf{l} \in \mathcal{I}^{(t+1)} \setminus \mathcal{I}^{(t)}} \|\Delta_{\mathbf{l}} u\|_{\mathcal{X}}. \tag{52}$$

The algorithm stops if this difference is smaller than  $\varepsilon$   $r$ -times in a row. The parameter  $r$  is introduced to avoid stopping too early, when the solution differs only slightly from one iteration to another while the next would result in a larger difference.

Regarding the overall computational cost, the adaptive approach has a computational overload. Although every full grid solution corresponding to  $\mathbf{l} \in \mathcal{A} \cup \mathcal{O}$  is computed during the algorithm, it is really only the solutions for  $\mathbf{l}$  close to the boundary of  $\mathcal{I}$  that are needed for computing the combination technique solution  $\mathbf{P}_{\mathcal{I}}^{\text{ct}} u$ . More precisely, only those solutions are used where  $c_{\mathbf{l}} \neq 0$ . In the case of nested spaces, the computational cost could be reduced by using the representation (47) of the sparse grid solution. This requires, however, the computation of the solutions in the hierarchical increment spaces, which we do not consider in this thesis.

## 6.2 Infinite-dimensional Parameter Space

An extended version of the above dimension-adaptive combination technique is described in this section. While we did not specify the quadrature rule in the previous section, we restrict ourselves to Gaussian quadrature rules here. The algorithm is modified such that the dimension of the parameter space can increase during the algorithm. Therefore, this version allows for problems where the truncation level of the Karhunen-Loeve expansion is not fixed beforehand and it can be seen as an algorithm in an infinite-dimensional stochastic parameter space, in other words  $m = \infty$ .

In the case of an infinite-dimensional stochastic domain, a few additional definitions need to be introduced. Computationally, we can only treat function evaluations with finitely many terms in the Karhunen-Loeve expansion. We denote by  $M$  during the algorithm the currently considered number of terms in the Karhunen-Loeve expansion and therefore the effective dimension of the parameter space. As the truncation of the Karhunen-Loeve expansion is equivalent to replacing the dependence of  $y_i$  for  $i > M$  by the constant 0, we use a similar concept to the anchored decomposition of a function in [62] and write as in Section 5.2

$$a_M(\mathbf{x}, y_1, \dots, y_M) = a(\mathbf{x}, y_1, \dots, y_M, 0, 0, \dots).$$

We notice that a replacement by the constant evaluation at 0 yields the same approximation in the parameter  $y_i$  as a quadrature rule with one point. As the random variables  $y_i$  are centered, a Gaussian quadrature with one point would estimate the integral over a constant function. We choose the number of quadrature points as  $N_{l_i}^{y_i} = 2^{l_i}$ , such that for  $l_i = 0$  the variable  $y_i$  is not regarded, while for  $l_i \geq 1$  a quadrature is applied.

Therefore, we call the stochastic dimension  $i$  active if  $l_i \geq 1$  and inactive if  $l_i = 0$ . Furthermore, to keep track of the stochastic variables that have been used, we say a stochastic dimension  $i$  is activated when the index  $\mathbf{e}_i$  is included in the index set  $\mathcal{O}$ . We note that this concept of activeness is different from the activeness of the index set  $\mathcal{A}$ .

The algorithm for the infinite-dimensional case proceeds in the same way as Algorithm 1 and is described in Algorithm 2. It starts with the parameter domain

being one-dimensional ( $M = 1$ ) and adds successively terms to the Karhunen-Loeve expansion. In addition to the search for new admissible neighbor indices to be added to  $\mathcal{A}$  for a refinement of the finite element solver or quadratures methods (line 10-18), we also need to check if a new variable  $y_{M+1}$  needs to be added to the consideration (line 19-27). Adding the parameter  $y_{M+1}$  corresponds to including the index  $\mathbf{e}_{M+1}$  in the active index set  $\mathcal{A}$ . As we need an admissible index set, this, however, can only be the case when the index  $\mathbf{j}$  selected in line 7 activates a new dimension.

Whenever a new variable  $y_{M+1}$  is taken into consideration, we need to extend the indices which were added in previous steps of the algorithm to be an  $(M + 1)$ -dimensional vector. We set the  $M + 1$ -th entry of the old indices to 0, as this is equivalent to a truncation after  $M$  terms. Finally,  $M$  is increased to  $M + 1$ . This procedure ensures that the dimension  $M$  is always one larger than the number of activated stochastic dimensions. The inclusion of  $\mathbf{e}_{M+1}$  in line 22 enables us to increase the stochastic dimension in the next step.

---

**Algorithm 2** Extended Dimension-adaptive Combination Technique

---

**Input:** Tolerance  $\varepsilon > 0$

**Output:** Index set  $\mathcal{I}$  and solution  $u_{\mathcal{I}}^{\text{ct}}$

```

1:  $\mathbf{l} = (0, 0)$ 
2:  $\text{act} = (0, 0, \dots)$ 
3:  $\mathcal{A} = \{\mathbf{l}\}, \mathcal{O} = \emptyset$ 
4: Compute local profit indicator  $\eta_{\mathbf{l}}$ 
5:  $\eta = \eta_{\mathbf{l}}$ 
6: while  $\eta > \varepsilon$  do
7:   Select index  $\mathbf{j} \in \mathcal{A}$  with largest profit  $\eta_{\mathbf{j}}$ 
8:    $\mathcal{A} = \mathcal{A} \setminus \{\mathbf{j}\}, \mathcal{O} = \mathcal{O} \cup \{\mathbf{j}\}$ 
9:    $\eta = \eta - \eta_{\mathbf{j}}$ 
10:  for  $k = 1, \dots, d$  do
11:     $\mathbf{l} = \mathbf{j} + \mathbf{e}_k$ 
12:    if  $\mathbf{l} - \mathbf{e}_i \in \mathcal{O} \quad \forall 1 \leq i \leq d$  then
13:       $\mathcal{A} = \mathcal{A} \cup \{\mathbf{l}\}$ 
14:      Compute full grid solution  $\mathbf{P}_{\mathbf{l}}u$ 
15:      Compute local profit indicator  $\eta_{\mathbf{l}}$ 
16:       $\eta = \eta + \eta_{\mathbf{l}}$ 
17:    end if
18:  end for
19:  if it exists  $n \in \{1 \dots, M\}$  s.t.  $\mathbf{j}_n > 0$  and  $\text{act}_n = 0$  then
20:     $\text{act}_n = 1$ 
21:     $\mathbf{l} = \mathbf{e}_{M+1}$ 
22:     $\mathcal{A} = \mathcal{A} \cup \{\mathbf{l}\}$ 
23:     $M = M + 1$ 
24:    Extend indices in  $\mathcal{A}, \mathcal{O}$ 
25:    Compute local profit indicator  $\eta_{\mathbf{l}}$ 

```

```

26:      $\eta = \eta + \eta_l$ 
27:   end if
28: end while
29: Combine solutions  $P_l u$  for  $l \in \mathcal{A} \cup \mathcal{O}$  according to (48) to obtain  $P_{\mathcal{I}}^{\text{ct}} u$  (48)
30: return  $\mathcal{I} = \mathcal{A} \cup \mathcal{O}$  and solution  $P_{\mathcal{I}}^{\text{ct}} u$ 

```

---

Looking at the extended dimension-adaptive combination technique from a theoretical perspective, the method iteratively refines one of three directions: either the spatial discretization, the quadrature or the truncation of the infinite-dimensional parameter domain: A finer discretization of the spatial domain is governed by the first entry of the index. If instead the quadrature method is refined, only one quadrature level among the stochastic dimensions that have already been activated is increased. The choice of this dimension depends on the profit indicator. The truncation level is raised when a new stochastic dimension is activated. Taking this perspective, we can choose an appropriate profit indicator.

### Profit Indicator

Next, we discuss how to choose the local profit indicator in the infinite-dimensional case. Again, we want to use the ratio of the benefit obtained by adding the index to  $\mathcal{O}$  and the cost required for the full grid computation. As before, the surplus  $\Delta_l u$  is used to measure how much an index contributes to the solution. But in contrast to the finite-dimensional case, we need to consider, in addition to the quadrature and spatial discretization levels, also the varying truncation level.

We recall that the first entry of  $l$  is the level for the spatial discretization and the  $l_{2:\infty} = (l_2, l_3, \dots)$  are the quadrature levels. We set  $k = \max\{i \in \mathbb{N} : l_i \geq 1\}$  for the truncation level. The benefit then is given by

$$\Delta_l u = (\Delta_{l_{2:\infty}}^y \otimes \Delta_{l_1}^x \otimes \Delta_k^T) u, \quad (53)$$

where

$$\Delta_l^y = \bigotimes_{\substack{i=1 \\ l_i \neq 0}}^{\infty} \Delta_{l_i}^{y_i} \quad \text{with} \quad \Delta_j^{y_i} = \begin{cases} Q_j^{(i)} - Q_{j-1}^{(i)} & \text{if } j > 1, \\ Q_1^{(i)} & \text{if } j = 1. \end{cases}$$

For the difference operator of the quadrature, we only take into account the stochastic dimensions which are active, meaning  $l_{i+1} \geq 1$ . For a simpler notation, we set  $\Delta_0^{y_i} = \text{Id}$  if the function is constant in  $y_i$ . Hence, we do not need to distinguish between active and inactive dimensions in the product of difference operators.

We focus in this part of the thesis on the computation of the expectation of the solution to the PDE. Thus,  $\Delta_l^x$  is the difference operator for the projection  $P$  of the PDE solution to the finite element space.

$$\Delta_l^x u = \begin{cases} P_l u - P_{l-1} u & \text{if } l > 0 \\ P_0 u & \text{if } l = 0. \end{cases}$$

For the computation of the second moment, we would need to define  $\Delta_j^x$  similar to (32).

The difference operator with respect to the truncation level is denoted here by

$$\Delta_k^T u = \begin{cases} u_k - u_{k-1} & \text{if } k > 1 \\ u_1 & \text{if } k = 1 \end{cases}$$

with  $u_k$  being the solution of the PDE when the diffusion coefficient  $a$  is replaced by the truncation of the Karhunen-Loeve expansion  $a_k$ . To have a well-defined operator  $\Delta_l$  in (53), we drop the dependence of  $y_j$  in  $u_k$  whenever  $l_j = 0$  for  $j < k$ .

For an easier implementation of the error contributions (53), we can make use of the fact that an inactive variable  $y_i$  corresponds to a quadrature level  $l_{i+1} = 0$ . We can reformulate (53) if  $l_{k+1} > 1$  as

$$\begin{aligned} \Delta_l u &= (\Delta_{l_{2:\infty}}^y \otimes \Delta_{l_1}^x \otimes \Delta_k^T) u \\ &= \left( \bigotimes_{i=1}^{k-1} \Delta_{l_{i+1}}^{y_i} \right) \otimes \left( \Delta_{l_{k+1}}^{y_k} \otimes \Delta_{l_1}^x \right) (u_k - u_{k-1}) \\ &= \left( \bigotimes_{i=1}^{k-1} \Delta_{l_{i+1}}^{y_i} \right) \otimes \left( (\Delta_{l_{k+1}}^{y_k} \otimes \Delta_{l_1}^x) u_k - \left( (Q_{l_{k+1}}^{(k)} \otimes \Delta_{l_1}^x) u_{k-1} - (Q_{l_{k+1}-1}^{(k)} \otimes \Delta_{l_1}^x) u_{k-1} \right) \right) \\ &= \left( \bigotimes_{i=1}^{k-1} \Delta_{l_{i+1}}^{y_i} \right) \otimes \left( (\Delta_{l_{k+1}}^{y_k} \otimes \Delta_{l_1}^x) u_k - (\Delta_{l_1}^x u_{k-1} - \Delta_{l_1}^x u_{k-1}) \right) \\ &= \left( \left( \bigotimes_{i=1}^k \Delta_{l_{i+1}}^{y_i} \right) \otimes \Delta_{l_1}^x \right) u_k, \end{aligned}$$

where we used in the second to last equation the fact that constant functions are integrated exactly by the quadrature rules.

In the case of  $l_{k+1} = 1$ , we notice that

$$Q_0^{(k)} \Delta_{l_1}^x (u_k - u_{k-1}) = 0.$$

Hence we can replace  $\Delta_1^{y_k}$  by  $\bar{\Delta}_1^{y_k} = Q_1^{(k)} - Q_0^{(k)}$  and use the same reasoning as above.

Using again the fact that constant functions are integrated exactly, (53) can be written as

$$\begin{aligned} \Delta_l u &= \left( \bigotimes_{\substack{i=1 \\ l_{i+1} \neq 0}}^k \Delta_{l_{i+1}}^{y_i} \otimes \Delta_{l_1}^x \right) u_k \\ &= \left( \bigotimes_{i=1}^k \bar{\Delta}_{l_{i+1}}^{y_i} \otimes \Delta_{l_1}^x \right) u_k \end{aligned} \tag{54}$$

with the modified first order difference operator

$$\bar{\Delta}_j^{y_i} = \begin{cases} Q_j^{(i)} - Q_{j-1}^{(i)} & \text{if } j > 0 \\ Q_0^{(i)} & \text{if } j = 0. \end{cases}$$

Hence, by (54) the local error contribution can be estimated with the same error estimator as in the finite-dimensional case and is applied to all variables that have been activated.

We also need to modify the cost function for the extended dimension-adaptive combination technique. As the number of terms in the Karhunen-Loeve truncation is not constant and could possibly grow to infinity, we need to incorporate the cost of evaluating the random field. We therefore use the cost function

$$c_{\mathbf{l}} = (N_{l_1}^x)^n \cdot \prod_{i=1}^{\infty} N_{l_{i+1}}^{y_i} \cdot k = (N_{l_1}^x)^n \cdot \prod_{i=1}^k N_{l_{i+1}}^{y_i} \cdot k$$

with the spatial dimension  $n$  and  $k = \max\{i: \mathbf{l}_i \geq 1\}$  being the maximal number of stochastic variables considered for  $\mathbf{l}$ . Here,  $N_{l_1}^x$  and  $N_{l_{i+1}}^{y_i}$  are again the number of degrees of freedom for the PDE solver and number of quadrature points, respectively. Another cost model that is not investigated here looks only at the number of active indices and uses  $\tilde{k} = |\{i \geq 2: \mathbf{l}_i \geq 1\}|$  instead of  $k$  (cf. [45], chapter 7).

Provided the error contribution and cost, we can define the local profit indicator as before by

$$\eta_{\mathbf{l}} = \frac{\|\Delta_{\mathbf{l}} u\|_{\mathcal{X}}}{c_{\mathbf{l}}}. \quad (55)$$

As the algorithm only looks for indices in the direct neighborhood of  $\mathcal{O}$ , it might happen that the algorithm stops too early. To avoid including a too small number of stochastic variables, which is the case when the error indicator is small for one of the stochastic parameters and more stochastic parameters should be considered, the algorithm can easily be modified to be more robust. Instead of considering in the index set  $\mathcal{A}$  only one index which increases the number of activated variables, a fixed number of indices corresponding to non-activated variables can be included into  $\mathcal{A}$ . By this, the algorithm can activate a stochastic dimension, even if the local error indicator is small for lower terms in the Karhunen-Loeve expansion. The number of considered stochastic variables is increased whenever the fixed number of non-activated variables has been reduced by the activation of one variable. Therefore, for the truncation level, possibly not one but a fixed number of forward neighbor indices are considered.

### 6.3 Numerical Examples

As a first example, we look at an integration problem in two dimensions to illustrate how the dimension-adaptive combination technique adjusts the index set  $\mathcal{I}$  to the applied numerical discretizations and the considered problem. In this case, solving the PDE on the spatial domain  $D$  is replaced by an integration problem for which we can apply one of the discussed quadrature rules.

We consider the integral

$$\mathbb{E} \left[ \int_0^1 u(x, y) dx \right] = \int_0^1 \int_0^1 u(x, y) dx dy$$

for different functions  $u$ . Furthermore, we use the trapezoidal rule for the variable  $x$ , while for the  $y$  variable, either a Gaussian quadrature or a trapezoidal rule as well is applied. We look at these two methods to demonstrate how the index set also depends on the applied numerical methods.

We use the above described profit indicator to select the next index and a global stopping criterion as in (51). The  $\mathcal{X}$ -norm, however, needs to be replaced by the absolute value for this integration problem. The resulting index sets for the functions

$$\begin{aligned} u_1(x, y) &= e^x + e^y + e^{xy}, \\ \text{and} \quad u_2(x, y) &= e^{-x^2} + 100e^{-y^2} \end{aligned}$$

are depicted in Figure 13.

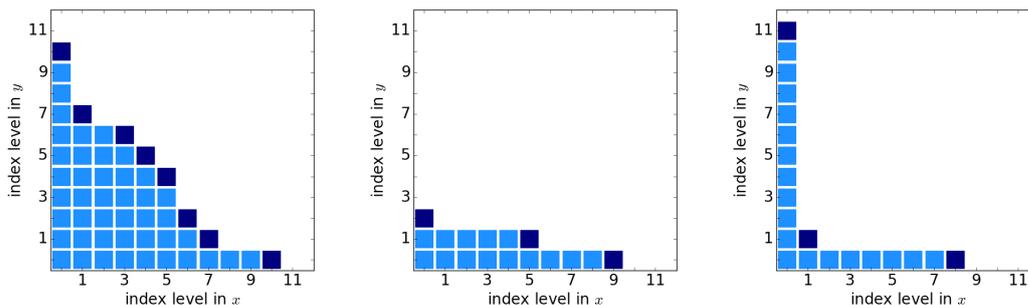


Figure 13: Index sets  $\mathcal{O}$  (light) and  $\mathcal{A}$  (dark) determined by the dimension-adaptive combination technique with the stopping criterion  $\eta \leq 10^{-9}$ . The first two panels show the index sets for  $u_1$  using a trapezoidal rule in  $y$  (left), and a Gaussian quadrature (middle). Right panel shows index sets for  $u_2$  using a trapezoidal rule for  $x$  and  $y$ .

First of all, it is observable that for the function  $u_1$  the adaptive approach provides a similar index set to the one expected by the a-priori analysis. Comparing the results for  $u_1$  with the two different methods in  $y$ , it can be expected that less quadrature points are needed for a Gaussian quadrature, as the convergence is faster. Indeed, the shape of the index set is close to an isosceles triangle with a trapezoidal rule, while for the Gaussian quadrature, less quadrature levels are used than for the quadrature in  $x$ .

In addition to the difference in the index sets due to the numerical methods applied, the constructed index set is also problem-dependent. To illustrate this, we look at a different integrand  $u_2$ , which is, similar to the function  $u_1$ , smooth. The function  $u_2$  can be separated into a sum of a function in  $x$  and a function in  $y$ . Therefore, the obtained index set does not have indices with medium cost

in both directions. Rather, the approach refines one direction while keeping the other fixed at the level of possible lowest work. In this case, the adaptive algorithm finds a better index set than we would get from the a-priori theory for functions with mixed regularity. Furthermore, a difference in the number of levels in the two directions is discernible, although we used the same numerical method for the integration in  $x$  and  $y$ . The maximal level is slightly higher for the variable  $y$ . This is due to the fact that the factors of the two terms are different. The dimension-adaptive algorithm resolves this anisotropic problem appropriately.

As a second example for the dimension-adaptive combination technique with a finite-dimensional parameter space, we look at a partial differential equation in one dimension which depends on only one random variable. The overall number of directions in this example is kept at two to be able to visualize the index sets.

For the example, we consider the model problem

$$-\operatorname{div}(a(x, y)\nabla u(x, y)) = f(x, y) \quad \text{in } D = (0, 1)$$

with  $y \sim \mathcal{U}(0, 1)$  and  $a(x, y) = 2xy + 1$  such that the  $a > 0$  for all  $y$  and apply the algorithm with

$$f(x, y) = 2\pi \sin(\pi y) (2y \sin(2\pi x) + 2\pi (2xy + 1) \cos(2\pi x))$$

such that the solution is smooth in  $x$  and  $y$  and known analytically:

$$u_{\text{exact}}(x, y) = \sin(\pi y) \cos(2\pi x).$$

As the quantity of interest, we are interested in the computation of the expectation  $\mathbb{E}[u(x, y)]$ . For that, we solve the PDE with piecewise polynomial elements, while for the stochastic approximation, either a trapezoidal or Gauss-Legendre quadrature is applied. Moreover, we evaluate the error of the computed result in the  $H^1(D)$ -norm. Therefore, we consider the  $H^1(D)$ -norm for the local error contribution estimation (50).

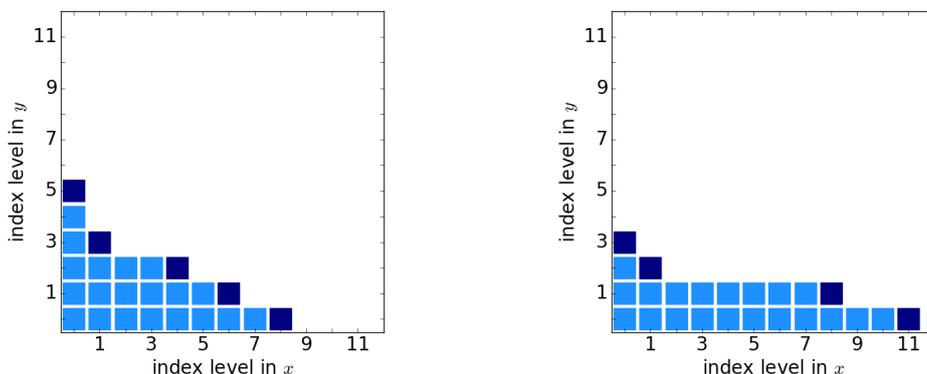


Figure 14: Index sets determined by the adaptive algorithm after 20 iterations with a trapezoidal rule (left) and a Gauss-Legendre quadrature (right). The dark blue represents the indices in  $\mathcal{A}$  and the light blue in  $\mathcal{O}$ .

In Figure 14, the index sets obtained after 20 iterations are visualized for the two cases. Similar as for the integration problem, we can detect differences in the index sets. The index set for the trapezoidal rule includes refinements of the PDE solver and only slightly less levels for the quadrature method. This can be expected, as the convergence rate of the PDE solver in the  $H^1(D)$ -norm is of order 1 and for the trapezoidal rule of order 2. Using the Gauss-Legendre quadrature corresponds to a spectral approach for the stochastic domain. This faster convergence of the Gaussian quadrature is identified by the algorithm and is reflected in the index set, which includes significantly higher levels in the spatial direction than for the quadrature.

In Figure 15, the convergence rates are shown with respect to the computational cost. Since the solution is smooth with respect to the combination of  $x$  and  $y$ , the a-priori analysis would provide a convergence rate for the error measured in the  $H^1(D)$ -norm of essentially  $\mathcal{O}(N^{-1})$ . This rate is obtained by the dimension-adaptive algorithm for both quadrature rules. Furthermore, we plot the error measured in the  $L^2(D)$ -norm. As the finite element method converges with the rate 2 in the  $L^2(D)$ -norm, we could expect to achieve a higher convergence rate. However, since the profit indicator is based on the  $H^1(D)$ -norm, the  $L^2$ -error can also only decrease with the rate  $\mathcal{O}(N^{-1})$ . Nevertheless, a higher convergence rate is in part observable for the Gaussian quadrature because the number of quadrature points is doubled when adding a higher quadrature level which increases the number of points more than needed.

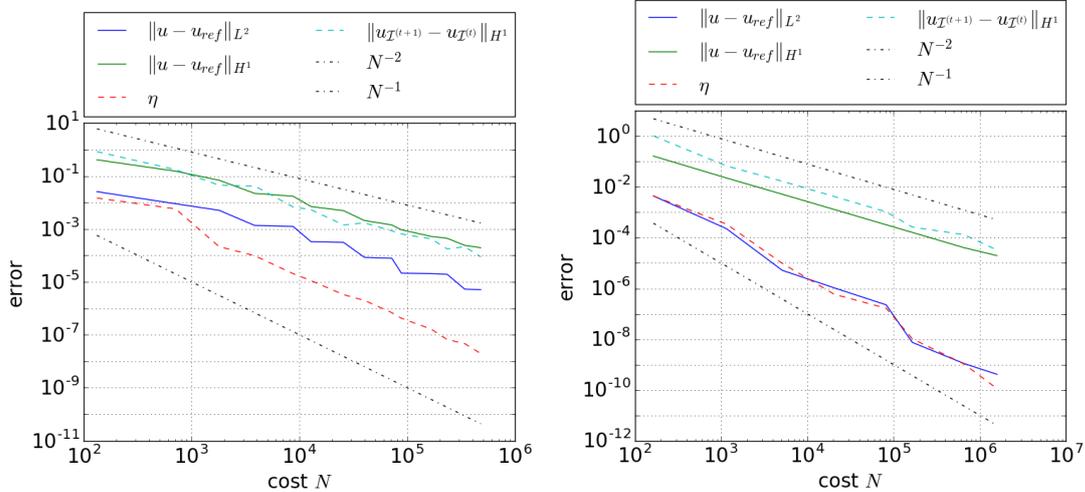


Figure 15: Convergence plots for the mean of the solution to the PDE. The left plot shows the convergence for the trapezoidal rule and the right for a Gauss-Legendre quadrature. The error is measured in the  $L^2(D)$ - and  $H^1(D)$ -norm. The dotted lines represent the rate  $N^{-1}$  and  $N^{-2}$  in both plots. Two different stopping criteria are shown.

Figure 15 also displays the different stopping criteria proposed above. The criterion (52), which evaluates the difference between the solutions of two iterations, captures properly the decrease of the  $H^1$ -error. Nevertheless, small oscillations

are visible, even though the parameter  $r$  was set to 8. Finally, the global profit indicator  $\eta$  (cf. (51)) is also shown. Rather than the  $H^1$ -error, the  $H^1$ -error per unit cost is estimated by this choice. For the finite-dimensional case, the choice of stopping criterion can therefore be chosen dependent on the situation, i.e., whether the error or profit should be smaller than a prescribed threshold.

Next, we apply the extended version of the dimension-adaptive combination technique. To this end, we consider an example where the diffusion coefficient of the PDE is given by a random field. We compute the Karhunen-Loeve expansion for a random field with mean  $\mathbb{E}[a(\mathbf{x})] = 8$  and a smooth covariance function

$$\text{cov}_a(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \exp\left(-\frac{1}{2} \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{\rho^2}\right), \quad (56)$$

where the correlation length is  $\rho = 0.2$  and the variance is chosen as  $\sigma^2$ . For the random field given in the Karhunen-Loeve expansion

$$a(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \psi_k(\mathbf{x}) y_k,$$

we assume that the random variables  $y_k$  are uniformly distributed in  $(-\sqrt{3}, \sqrt{3})$  to ensure coercivity. A realization of the field and the decay of the coefficients in the Karhunen-Loeve expansion is shown in Figure 16.

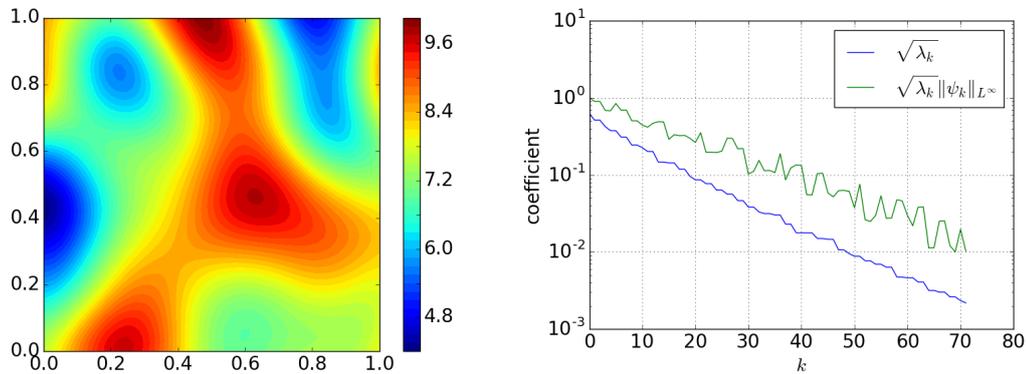


Figure 16: Left: realization of the random field with covariance function as in (56). Right: decay of coefficients in the Karhunen-Loeve expansion.

We compute for this choice of diffusion coefficient the expectation  $\mathbb{E}[u(x, \mathbf{y})]$  where  $u(x, \mathbf{y})$  solves for a.e.  $\mathbf{y} \in \Gamma$

$$\begin{aligned} -\operatorname{div}(a(x, \mathbf{y}) \nabla u(x, \mathbf{y})) &= f(x) && \text{in } D = (0, 1)^2, \\ u(x, \mathbf{y}) &= 0 && \text{on } \partial D \end{aligned}$$

with  $f(x) = 100$ . Moreover, we apply the adaptive algorithm using a profit indicator (55) based on the  $L^2(D)$ -norm, i.e., the local error contribution is given by  $\|\Delta_I u\|_{L^2(D)}$ .

The convergence of the algorithm in terms of the number of iterations and the computational cost is shown together with the two stopping criteria (51) and

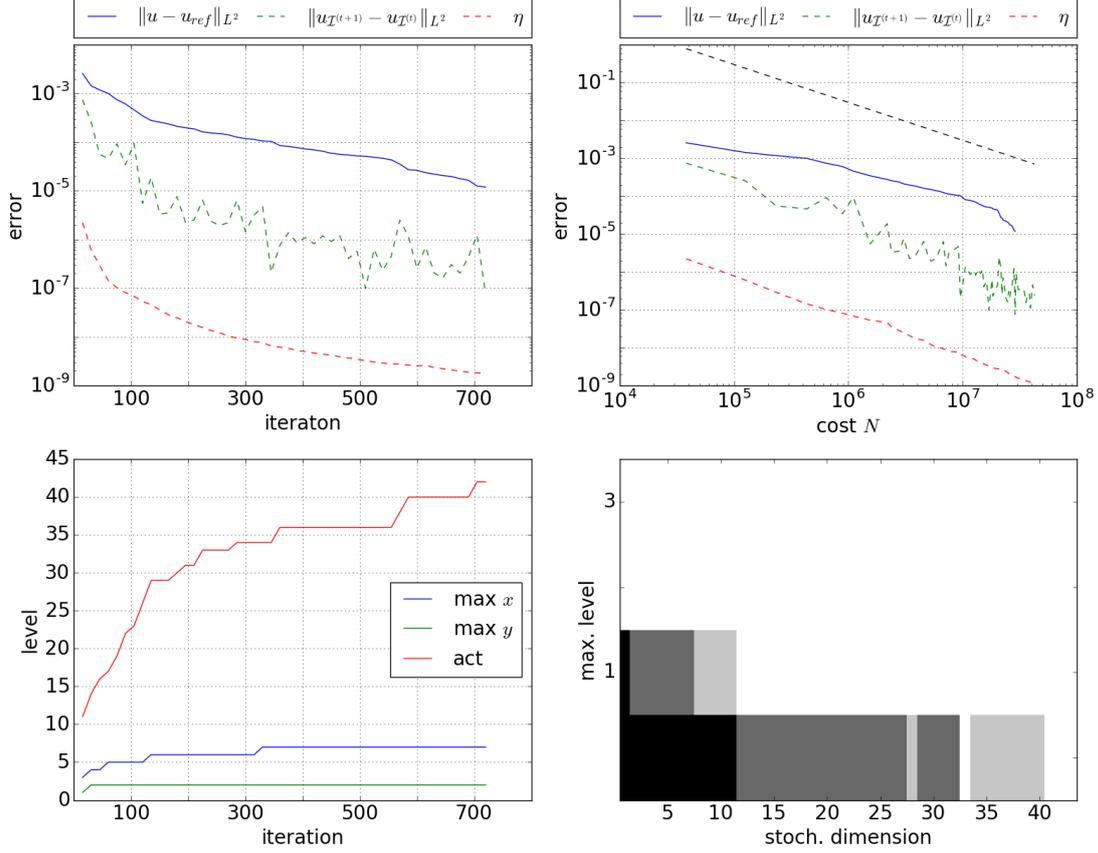


Figure 17: Top row: convergence plot of the dimension-adaptive algorithm plotted against the iteration number (left) and computational cost (right). The black dashed line corresponds to the rate  $\mathcal{O}(N^{-1})$ . Bottom left: maximal levels in the index set  $\mathcal{I}$  for the spatial refinement, for all random variables and for the truncation level. Bottom right: quadrature level in  $\mathcal{O}$  for each activated random variable at three different iteration steps (black: it. 30, dark: it. 300, light: it. 720).

(52) in Figure 17. The error is computed based on a reference solution obtained with additional 150 iterations, as the exact solution is unknown. We observe that the error decreases with increasing number of iterations. For the first iterations, the increase in the computational cost is high when many stochastic dimensions are activated and many indices are added to  $\mathcal{A}$ , so that the error only decreases slowly at first with respect to the computational cost. After that, the algorithm achieves a convergence rate of almost  $\mathcal{O}(N^{-1})$ . However, with increasing activated stochastic variables, the number of indices in the neighborhood set  $\mathcal{A}$  which correspond to a refinement of the quadrature becomes large. Many iterations are therefore needed to reduce the quadrature error, even though these iterations are not expensive.

The obtained convergence results are consistent with the expected behavior from the a-priori analysis. Since the convergence rate for solving the PDE is of order 1, this is the maximal rate which can be expected by the combination technique. We see that this rate is almost obtained. The adaptive approach with the

Gaussian quadrature yields, therefore, a better rate than the often applied Monte Carlo quadrature, which limits the convergence rate to  $-1/2$ .

The bottom left plot in Figure 17 illustrates the evolution of the index set. It shows how the maximal level considered for each of the three refinement directions (spatial, stochastic, truncation) increases during the algorithm, whereby for the quadrature the maximal level in all stochastic dimensions is considered. As expected, a finer discretization of the spatial domain is needed than for the stochastic parameters. Since the Gaussian quadrature converges exponentially, low discretization levels for the quadrature are sufficient, which is reflected by 2 being the highest level in the index set  $\mathcal{A}$ .

Moreover, the number of activated stochastic dimensions increases step by step during the algorithm. We observe that while most of the stochastic variables have been included into the consideration during the first iterations, additional stochastic variables are added sporadically whenever the other directions are approximated fine enough.

The refinement of each stochastic dimension is shown in the lower right panel of Figure 17. It illustrates the quadrature levels for each stochastic variable at different points during the algorithm. While the first stochastic dimensions are refined to a higher level, only the lowest level quadrature method is needed for the higher stochastic dimensions. We observe that the algorithm adapts to the anisotropy of the solution caused by the decay in the Karhunen-Loeve expansion. It only refines those stochastic dimensions that contribute the most to the solution, similar to the construction of the anisotropic product quadrature.

The algorithm activates during the 720 iterations 40 stochastic dimensions which capture the uncertainty to an acceptable extent. The number of stochastic variables included depend on the decay of the coefficients in the Karhunen-Loeve expansion which in turn depend on the smoothness of the random field and correlation length. Additional numerical examples show that random fields with less smooth covariance kernels, require more stochastic variables, as the decay of the coefficients is faster. In contrast, the algorithm activates less stochastic variables when the correlation length is longer, matching the expected behavior. Since apart from this difference the properties of the sparse grid are similar, these examples are not presented here.

As a last example in this section, we apply the extended version of the dimension-adaptive combination technique to a problem which we have not considered in the a-priori analysis. To this end, we consider a problem in one dimension with a diffusion coefficient which is a lognormal field. This choice models the diffusion coefficient better in some applications, for example in groundwater flow problems, which will be discussed in the next section.

The diffusion coefficient is written in this case as

$$a(x, \mathbf{y}) = \exp(b(x, \mathbf{y})),$$

where  $b(x, \mathbf{y})$  is a Gaussian field given in the Karhunen-Loeve expansion. Hence, also  $a$  depends on infinitely many parameters. Moreover, a lognormal field ensures

that  $a(\mathbf{x}, \omega)$  is positive even if normal distributed random variables are considered in the Karhunen-Loeve expansion.

We apply the dimension-adaptive combination technique in order to compute the expectation to the one-dimensional PDE, where the forcing term is chosen as  $f(x) = 1$ , and assume the field  $b$  has zero mean and a Matérn covariance function with smoothness parameter  $\nu = 2.5$  and correlation length  $\rho = 0.2$ . The decay of the coefficients  $\{\sqrt{\lambda_k}\}_k$  are shown in Figure 19.

As the dependence of the random field on the stochastic variables is different, we cannot directly compare this example with the previous one. Nevertheless, from the decay of the coefficients, we also expect that the first stochastic variables need a higher quadrature rule than those with a smaller coefficient. Indeed, the algorithm adapts to the anisotropy in the random variables and refines the dependence on the stochastic variables to a different extent, as visible in Figure 18.

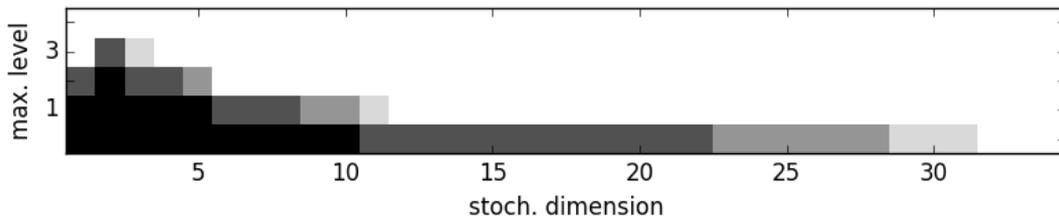


Figure 18: The maximal level in  $\mathcal{O}$  for each stochastic dimension at four different iteration steps (black: it. 160, dark: it. 1280, medium: it. 2560, light: it. 4000).

Overall, we observe a similar behavior of the error as in the previous example. In the top row of Figure 19, the convergence of the algorithm in terms of the number of iterations and the computational cost is shown. Again, a reference solution is used for the error computation. We observe that also for the lognormal case the convergence rate is determined by the finite element solver. Although we consider here a one-dimensional spatial domain, we obtain a convergence rate essentially of order 1, because we chose the profit indicator based on the  $H^1(D)$ -norm this time.

In this and also in the previous example, we observe a different behavior of the global error indicator  $\eta$  than in the case where the stochastic dimension was fixed. While  $\eta$  behaved as the error per unit cost for the finite-dimensional case, here the global error indicator decreases more slowly. This might be due to the fact that the number of neighboring indices in  $\mathcal{A}$  which correspond to a refinement of the quadrature grows when more stochastic variables are activated. Hence, more terms are considered in the sum (51).

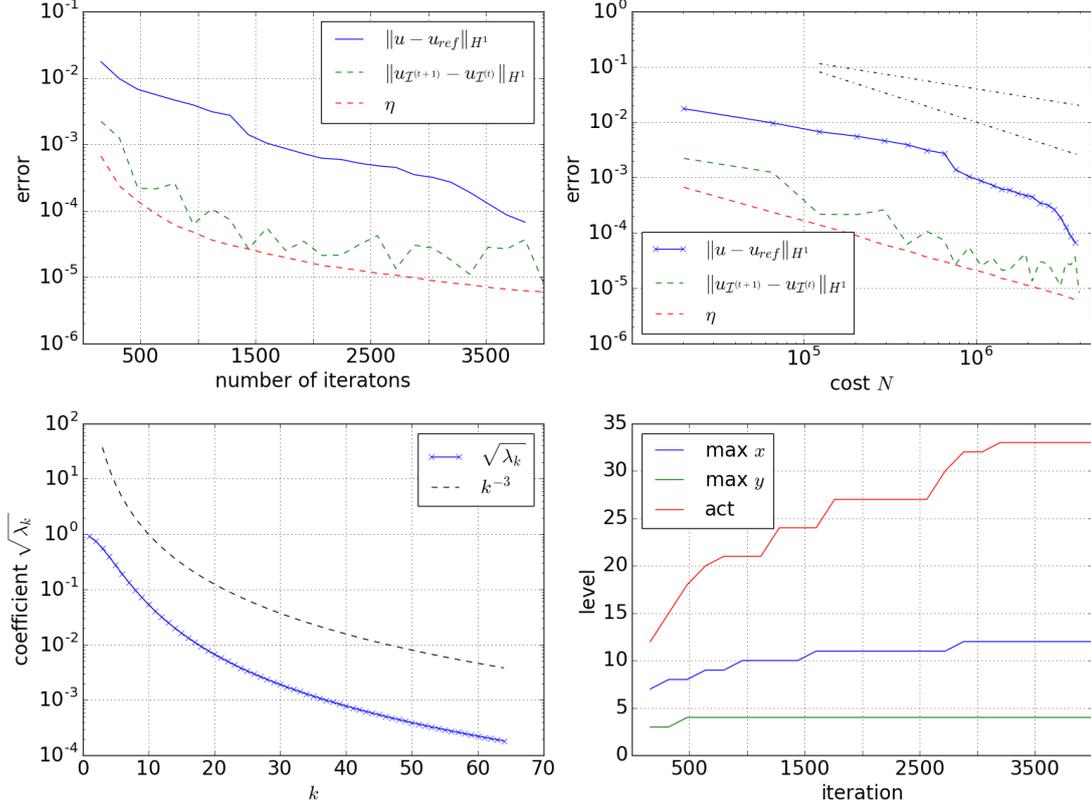


Figure 19: Top row: convergence plot of the dimension-adaptive algorithm applied to a lognormal problem plotted against the iteration number (left) and the computational cost (right), where the dashed lines correspond to rates  $N^{-0.5}$  and  $N^{-1}$ . Bottom left: decay of coefficients in Karhunen-Loeve. Bottom right: maximal levels in the index set  $\mathcal{I}$  for the spatial refinement, for all random variables, and for the truncation level.

Finally, we illustrate the interplay of the truncation, quadrature and finite element method in Figure 20 and plot the index set  $\mathcal{O}$  for one iteration step. We consider the maximal level over all quadrature rules as one parameter for the stochastic approximation. As expected, we observe that the shape is similar to a tetrahedron. For a high resolution in one direction, the other directions are resolved with a low discretization level. However, the shape is not exactly a tetrahedron, as the quadrature is not independent of the truncation level.

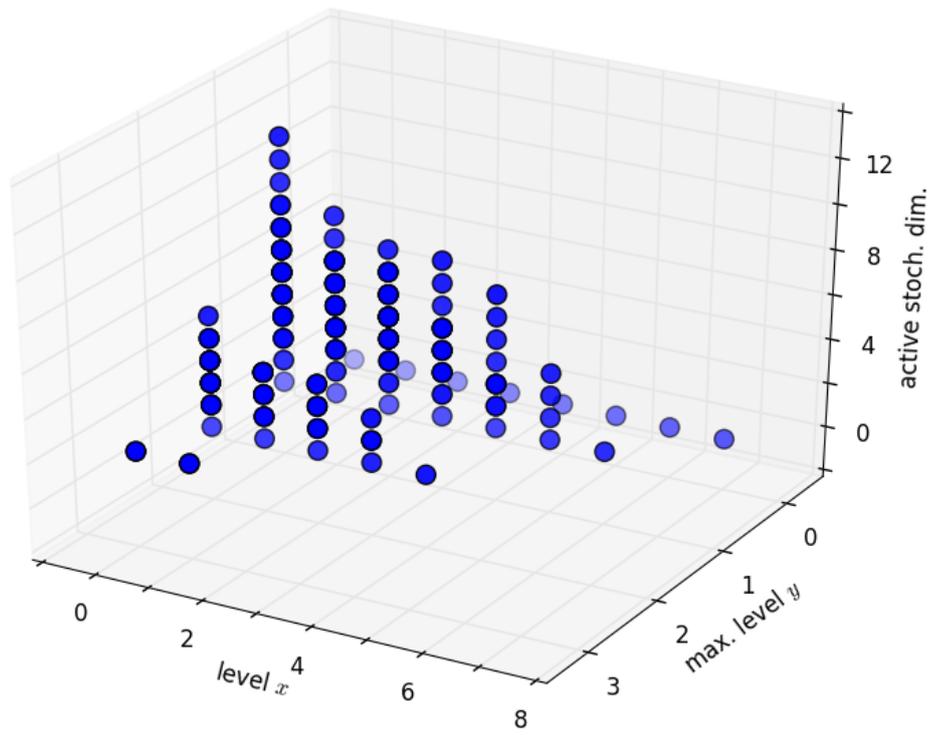


Figure 20: Visualization of index set  $\mathcal{O}$  after 200 iterations of the algorithm.

## 7 Application to Darcy Flow Problem

As an application of the dimension-adaptive combination technique, we will consider the Darcy problem, a problem in uncertainty quantification that arises in the study of groundwater flows. There uncertainty is introduced to describe a heterogeneity of the medium or to circumvent a lack of knowledge, as the physical properties of the medium are often unknown or only a small number or noisy measurements are available. The study of subsurface flows is relevant, for example, in water management, environmental and energy problems (see, e.g., [63]), where it is important to understand the uncertainty and its propagation through a flow model.

In the Darcy problem the flow of a single-phase fluid in a porous medium is described by Darcy's law together with the law of mass conservation. These laws are formulated in the following two equations:

$$\begin{aligned} a^{-1}(\mathbf{x})\mathbf{u}(\mathbf{x}) &= -\nabla p(\mathbf{x}) && \text{in } D, \\ \operatorname{div} \mathbf{u}(\mathbf{x}) &= f(\mathbf{x}) && \text{in } D, \end{aligned} \tag{57}$$

where  $D \subset \mathbb{R}^2$  is the spatial domain.

Here,  $\mathbf{u}$  describes the Darcy flux, the volume of fluid transported per total area and time. The coefficient  $a(\mathbf{x})$  describes the hydraulic conductivity which measures the transmissibility of the fluid through the porous medium, which in turn depends on the intrinsic permeability, a property of the medium, and the dynamic viscosity of the fluid. The function  $p$  is the hydraulic head, describing the fluid pressure, and the right hand side  $f$  accounts for a possible source term.

In this context, we assume that the spatial domain is rectangular. In addition to (57), boundary conditions need to be specified which can be either Dirichlet or Neumann boundary conditions. We distinguish the part of the boundary where Dirichlet and Neumann boundary conditions are provided by  $\partial D^D$  and  $\partial D^N$ .

$$p = p_0 \quad \text{on } \partial D^D \quad \text{and} \quad -\nabla p \cdot \mathbf{n} = g \quad \text{on } \partial D^N.$$

As mentioned above, the physical properties of the medium are subject to uncertainty and the hydraulic conductivity is described by a random field  $a(\mathbf{x}, \omega)$ . As a consequence, the Darcy flux  $\mathbf{u} = \mathbf{u}(\mathbf{x}, \omega)$  and pressure  $p = p(\mathbf{x}, \omega)$  are themselves random functions. The pair  $(\mathbf{u}, p)(\mathbf{x}, \omega)$  is a solution to the Darcy problem if (57) is satisfied for a.e.  $\omega \in \Omega$ .

The properties of the random field depend on the material of the sediment. The variations of the field are usually large and encompass several orders of magnitude. Therefore, the hydraulic conductivity is often modeled as a lognormal random field. An example of the permeability field which determines the hydraulic conductivity is shown in Figure 21. It is included in the SPE10 data set which is provided for reservoir modeling by the Society of Petroleum Engineers ([64]) and which is widely used in the study of groundwater flow (see, for example, [65, 66, 67]). This data set consists of a realization of a geostatistical

model describing the permeability of the Tarbert formation, which is located in the North Sea. It describes an area of  $1200 \times 2200$  feet where the soil is composed of sandstone, siltstone, and shales.

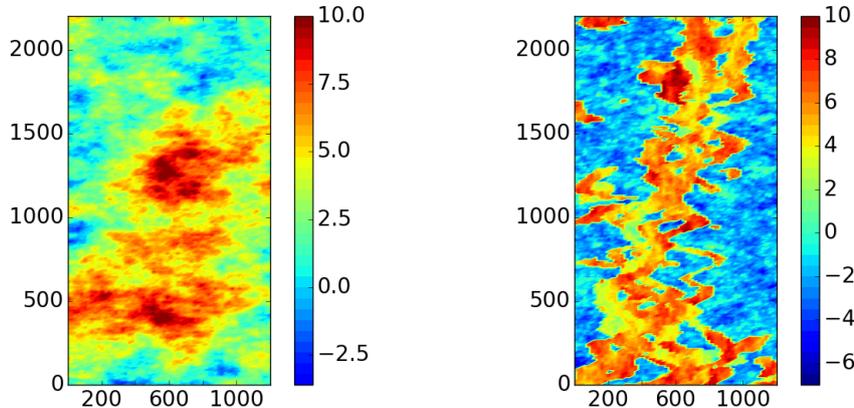


Figure 21: Logarithm of horizontal permeability field of two layers in the SPE10 data set in millidarcy. Left: Layer 10 composed of sandstone siltstone and shales. Right: Layer 50 which is part of the fluvial Ness formation.

This data set also includes sediment layers which represent the fluvial Ness formation (see right panel in Figure 21). The permeability in this formation is qualitatively different as the formation consists of sediment layers shaped by running water. Hence, areas of high-permeable sandstone channels with a long correlation length and low-permeable areas of shales and coal with short correlation lengths exist. The field therefore appears to follow rather a bi-modal lognormal distribution than a lognormal distribution ([66]).

For the application of the dimension-adaptive algorithm, we restrict ourselves to the case of a lognormal permeability field as shown in the left panel of Figure 21 corresponding to shallow-marine sediment layers. In this case, we can model the hydraulic conductivity as a random field where the logarithm is a Gaussian field, characterized by its mean and covariance function. Thus, we assume that  $a(\mathbf{x}, \omega) = \exp(b(\mathbf{x}, \omega))$ , where  $b(\mathbf{x}, \omega)$  is a Gaussian field. Like before, we expand the random field  $b(\mathbf{x}, \omega)$  into a Karhunen-Loeve expansion and parametrize the stochastic dependence, such that

$$a(\mathbf{x}, \mathbf{y}) = \exp \left( \sum_{k=1}^{\infty} \sqrt{\lambda_k} \psi_k(\mathbf{x}) y_k \right).$$

This leads to the parametrized version of (57). The solution  $(\mathbf{u}(\mathbf{x}, \mathbf{y}), p(\mathbf{x}, \mathbf{y}))$  solves for fixed  $\mathbf{y} \in \Gamma$

$$\begin{aligned} \operatorname{div} \mathbf{u}(\mathbf{x}, \mathbf{y}) &= f(\mathbf{x}) && \text{in } D \\ a^{-1}(\mathbf{x}, \mathbf{y}) \mathbf{u}(\mathbf{x}, \mathbf{y}) &= -\nabla p(\mathbf{x}, \mathbf{y}) && \text{in } D \\ &+ \text{boundary conditions.} \end{aligned} \tag{58}$$

This system can be rewritten as the second order elliptic PDE

$$-\operatorname{div}(a(\mathbf{x}, \mathbf{y}) \nabla p(\mathbf{x}, \mathbf{y})) = f \quad \text{in } D,$$

which is the model problem considered before and for which we applied the dimension-adaptive algorithm in Section 6.3. However, here we consider the mixed formulation which is better suited for groundwater flow problems, as also the flux  $\mathbf{u}$  is of interest. The mixed formulation has the advantage that it solves for  $(\mathbf{u}, p)$  simultaneously and preserves mass locally.

In recent years, multilevel and sparse grid approaches have been applied for the computations of quantities of interest for the Darcy problem, either in mixed formulation or as elliptic PDE. In [68], a multilevel Monte Carlo method is considered which balances solving the mixed formulation with the number of Monte Carlo samples. A sparse grid approach for the stochastic parameters after a truncation of the Karhunen-Loeve expansion has been used in [69]. Furthermore, an adaptive algorithm for the stochastic parameter set has been applied to the elliptic PDE in [61].

Here, we apply the dimension-adaptive algorithm from Section 6, taking into account the approximation error of PDE solving, stochastic approximation, and restriction to finitely many random variables. We compute in particular the expectations  $\mathbb{E}[\mathbf{u}(\mathbf{x})]$  and  $\mathbb{E}[p(\mathbf{x})]$  as quantities of interest. For that, the algorithm from Section 6.2 is modified, such that the mixed formulation (58) is used. Since the algorithm uses standard finite element code for solving the PDE for a point  $\mathbf{y} \in \Gamma$ , the algorithm can easily be adjusted to solve the mixed formulation instead. In the next section, we therefore discuss the finite element method used for solving the mixed problem for a fixed  $\mathbf{y} \in \Gamma$ .

## 7.1 Mixed Finite Elements

In order to solve (58) numerically, a mixed finite element method is applied. For that, we consider the weak formulation of the mixed formulation. For  $\mathbf{u}$  we define the function space  $H(\operatorname{div}, D) = \{\mathbf{v} \in L^2(D)^2 : \operatorname{div} \mathbf{v} \in L^2(D)\}$  together with the norm  $\|\mathbf{v}\|_{H(\operatorname{div})} := (\|\mathbf{v}\|_{L^2}^2 + \|\operatorname{div} \mathbf{v}\|_{L^2}^2)^{1/2}$ . Then, for  $\mathbf{y} \in \Gamma$ , the weak formulation reads: Find  $(\mathbf{u}, p) \in H(\operatorname{div}, D) \times L^2(D)$  such that

$$\begin{aligned} r(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) &= \int_{\partial D} p_0 \mathbf{n} \cdot \mathbf{v} \, dx & \forall \mathbf{v} \in H(\operatorname{div}, D) \\ b(\mathbf{u}, q) &= - \int_D f q \, dx & \forall q \in L^2(D) \end{aligned} \tag{59}$$

with the bilinear forms  $r : H(\operatorname{div}, D) \times H(\operatorname{div}, D) \rightarrow \mathbb{R}$  and  $b : H(\operatorname{div}, D) \times L^2(D) \rightarrow \mathbb{R}$  defined for all  $\mathbf{v}_1, \mathbf{v}_2 \in H(\operatorname{div}, D)$  and  $q \in L^2(D)$  by

$$\begin{aligned} r(\mathbf{v}_1, \mathbf{v}_2) &:= \int_D a^{-1} \mathbf{v}_1 \cdot \mathbf{v}_2 \, dx \\ b(\mathbf{v}_1, q) &:= - \int_D \operatorname{div}(\mathbf{v}_1) q \, dx. \end{aligned}$$

For simplicity, we assumed zero boundary conditions on  $D^N$  in the weak formulation, but note that the boundary conditions of the type  $-\nabla p \cdot \mathbf{n} = g$  must be integrated into the function space  $H(\operatorname{div}, D)$ , while the condition  $p = p_0$  on  $\partial D^D$  is natural and is included in the weak formulation. The existence of a unique solution to the saddle point problem (59) is guaranteed for lognormal random fields with a covariance function of  $\log(a)$  that is regular enough, including Matérn covariance functions (for details see [68] and references therein).

An approximate solution to the saddle point problem is found by solving the discrete version of (59). There the spaces  $H(\operatorname{div}, D)$  and  $L^2(D)$  are replaced by  $\mathcal{V}_h \subset H(\operatorname{div}, D)$  and  $\mathcal{W}_h \subset L^2(D)$  respectively.

As finite element spaces we choose the pair of lowest order Raviart-Thomas elements  $\mathcal{V}_h = RT_1 = \{\mathbf{v} \in L^2(D)^2: \mathbf{v}|_K = \boldsymbol{\alpha}_k + \beta_k \mathbf{x}, \boldsymbol{\alpha}_K \in \mathbb{R}^2, \beta_k \in \mathbb{R} \forall K \in \mathcal{T}_h$  and  $\mathbf{v} \cdot \mathbf{n}$  is constant across interfaces} and piecewise constant pressures  $\mathcal{W}_h = \{q \in L^2(D): q|_K \in \mathcal{P}_0 \forall K \in \mathcal{T}_h\}$  for a triangulation  $\mathcal{T}_h$  of  $D$ .

This pair is chosen because it is stable in the sense that it satisfies a discrete inf-sup condition and the bilinear form  $r$  is coercive on the set  $\{\mathbf{v}_h \in \mathcal{V}_h: (\mathbf{v}_h, q_h) = 0 \forall q_h \in \mathcal{W}_h\}$ . The pair of lowest order of Raviart-Thomas elements and discontinuous pressure exhibit a convergence rate for  $\|\mathbf{u}_h - \mathbf{u}\|_{H(\operatorname{div}, D)} + \|q_h - q\|_{L^2(D)}$  of order 1 at most. Under certain conditions higher convergence rates can be shown for the pressure variable [68, 70].

## 7.2 Numerical Examples

We will consider the flow problem on the unit square  $D = (0, 1)^2$  with  $f = 0$ . The boundary conditions are specified by

$$\begin{aligned} p &= 0 \text{ on } \partial D \cap \{x_1 = 1\}, \\ p &= 1 \text{ on } \partial D \cap \{x_1 = 0\}, \\ \nabla p \cdot \mathbf{n} &= 0 \text{ on } \partial D \cap \{x_2 = 0 \text{ or } x_2 = 1\}. \end{aligned}$$

The algorithm of Section 6.2 is applied to estimate  $\mathbb{E}[\mathbf{u}]$  and  $\mathbb{E}[p]$ . Since the finite element method has changed, the indicator, which is used to determine the direction the algorithm refines next, needs to be adjusted. We need to modify the local estimator for the error contribution in order to measure the benefit  $\|\Delta_{\mathbf{l}}(\mathbf{u}, p)\|$  of adding the index  $\mathbf{l}$  to the index set. To this end, we use the norm

$$\|(\mathbf{v}, q)\| = \|\mathbf{v}\|_{H(\operatorname{div}, D)} + \|q\|_{L^2(D)}.$$

for the pair of functions  $(\mathbf{v}, q) \in H(\operatorname{div}, D) \times L^2(D)$ . Hence, both components have an influence on the adaptive refinement. Corresponding to (55), the indicator then is chosen as

$$\eta_{\mathbf{l}} = \frac{\|\Delta_{\mathbf{l}}(\mathbf{u}, p)\|}{c_{\mathbf{l}}},$$

where  $c_{\mathbf{l}}$  denotes the cost.

For the hydraulic conductivity, we consider different lognormal random fields. In all cases we compute the Karhunen-Loeve expansion of  $\log(a(\mathbf{x}, \omega))$ . First, we look at two choices for the covariance function of the Gaussian field  $\log(a(\mathbf{x}, \omega))$ . We assume Matérn covariance functions, but with different smoothness parameter  $\nu$ . On the one hand, we consider a rather smooth random field  $b(\mathbf{x}, \omega)$  with  $\nu = 2.5$ , i.e., the covariance function is given by

$$\text{cov}(r) = \sigma^2 \left( 1 + \frac{\sqrt{5}r}{\rho} + \frac{5}{3} \frac{r^2}{\rho^2} \right) \exp \left( -\frac{\sqrt{5}r}{\rho} \right)$$

with  $r = \|\mathbf{x}_1 - \mathbf{x}_2\|_2$ , correlation length  $\rho = 0.15$  and  $\sigma^2 = 5$ . On the other hand, a rougher field is used, whereby the covariance function is chosen to be an exponential covariance function ( $\nu = 0.5$ ) with correlation length  $\rho = 0.25$  and  $\sigma^2 = 8$ . The mean of the field  $b$  is chosen as  $\mathbb{E}[b(\mathbf{x})] = 3$ .

For both choices a realization of the field  $b$  is shown in Figure 22, where it can be seen that they include values of the same magnitude as the data set in Figure 21.

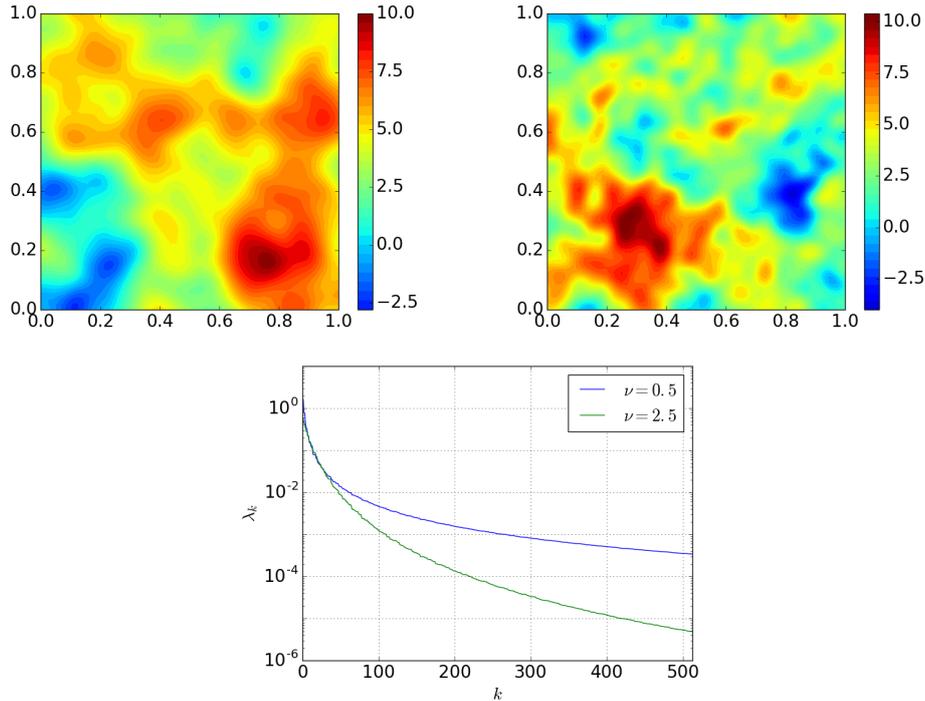


Figure 22: Top row: realizations of random fields with Matérn covariance functions. Left: a relatively smooth field with smoothness parameter  $\nu = 2.5$ . Right: rougher random field with  $\nu = 0.5$ . Bottom: decay of eigenvalues.

Figure 23 shows the behavior of the error for the flux and pressure measured in comparison with a reference solution obtained with 300 additional iterations of the algorithm. The algorithm activates for the first field 73 stochastic variables and 93 for the second. As the second random field is rougher, the decay of the coefficient is slower and thus more variables need to be included.

Apart from this difference, we can observe a similar behavior for both random fields. The error for the flux and pressure both decrease, whereby a larger error is observable for the flux. Since we use for the indicator the sum  $\|\mathbf{u}\|_{H(\text{div})} + \|p\|_{L^2(D)}$ , the error for  $\mathbf{u}$  and  $p$  do not decrease independently from each other. With respect to the computational cost we observe a convergence rate of approximately  $\frac{1}{2}$ . The convergence is limited to this rate, as the mixed finite element method on a two-dimensional domain does not converge faster. Considering this, the faster convergence by the application of a product of Gaussian quadratures is not exploited. A Monte Carlo method would lead to a similar convergence rate and would simplify the algorithm due to its independence of the dimension. However, we can expect to achieve higher convergence rates if the spatial discretization exhibit a faster rate with the discussed dimension-adaptive algorithm, which is not possible using a Monte Carlo method.

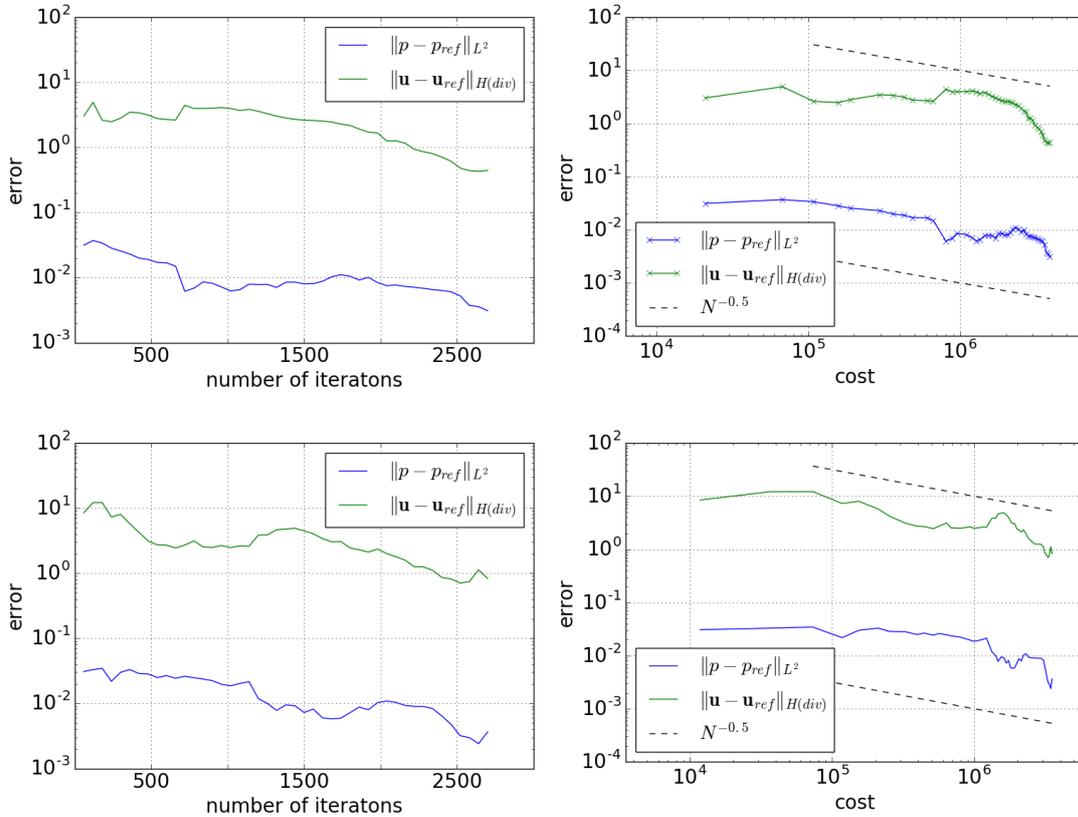


Figure 23: Convergence of dimension-adaptive combination technique applied to the Darcy flow problem. Top row displays errors for a lognormal field with parameter  $\nu = 2.5$ , bottom row for  $\nu = 0.5$ .

In a last example, we incorporate the points of the data set in the description of the random field. Instead of considering a random field that is structurally the same as the SPE10 data set, but does not include the data, we use for the next example the given values to construct the random field.

We assume that the data points describe the permeability field on the points exactly and want to model the field on a finer scale than provided by the data

points. For that, we consider a random field on the finer scale that interpolates the data points, but provides a random description of the field in between. This method of interpolating given data points is sometimes referred to as kriging.

Hence, the field under consideration arises from a different modeling problem than before. Instead of looking at general permeability fields, we focus here on a field that describes a particular sediment layer.

In order to obtain the random field described by a set of random variables, we apply a code provided by Paul Constantine [71] which computes the Karhunen-Loeve expansion and incorporates the available data points. However, only certain covariance functions for the field are available. Since the logarithm of the random field in the entire domain does not seem to be based on a Gaussian covariance function, we only consider a small subset of the data in the SPE10 data set of the first horizontal layer for our application. We want to compute the solution of the Darcy flow problem for this domain with a finer spatial resolution than in the previous examples.

For that, we look at a rectangular shaped domain of  $55 \text{ m} \times 27 \text{ m}$  which includes  $9 \times 9$  data points and scale the domain to  $D = (0, 1)^2$ . We find that the given data is best captured with  $\log(a(\mathbf{x}, \omega))$  being a Gaussian field with covariance function

$$\text{cov}(\mathbf{x}_1, \mathbf{x}_2) = 10 \exp\left(-\frac{1}{2} \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{0.01}\right).$$

Figure 24 shows the data points and a realization of the Gaussian field  $\log(a(\mathbf{x}, \omega))$  computed based on these points. For this random field, we look at the same PDE problem as before.

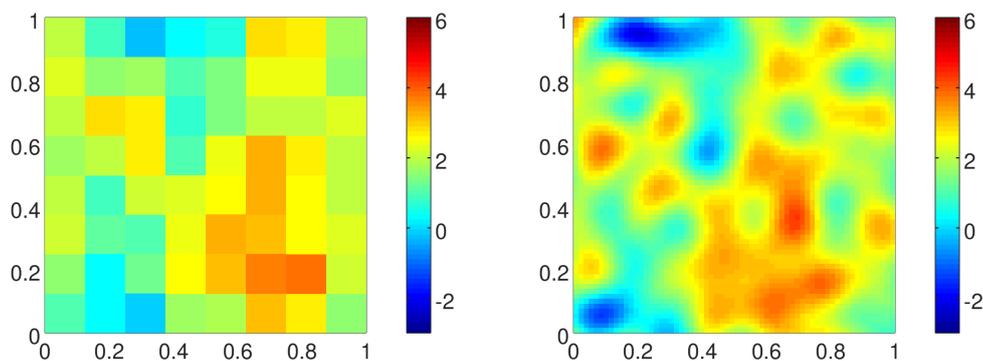


Figure 24: Left: logarithm of permeability field data points provided by SPE10 data set on a small part of the domain of the first layer. Right: realization of  $\log(a(\mathbf{x}, \omega))$  constructed to fit the data points.

The solution obtained by applying the dimension-adaptive combination technique is shown in Figure 25. The pressure and flux components are shown separately. As the mean of the field is not constant anymore, but rather based on the given data points, the expectation of flow and pressure reflects the different hydraulic conductivity in the domain.

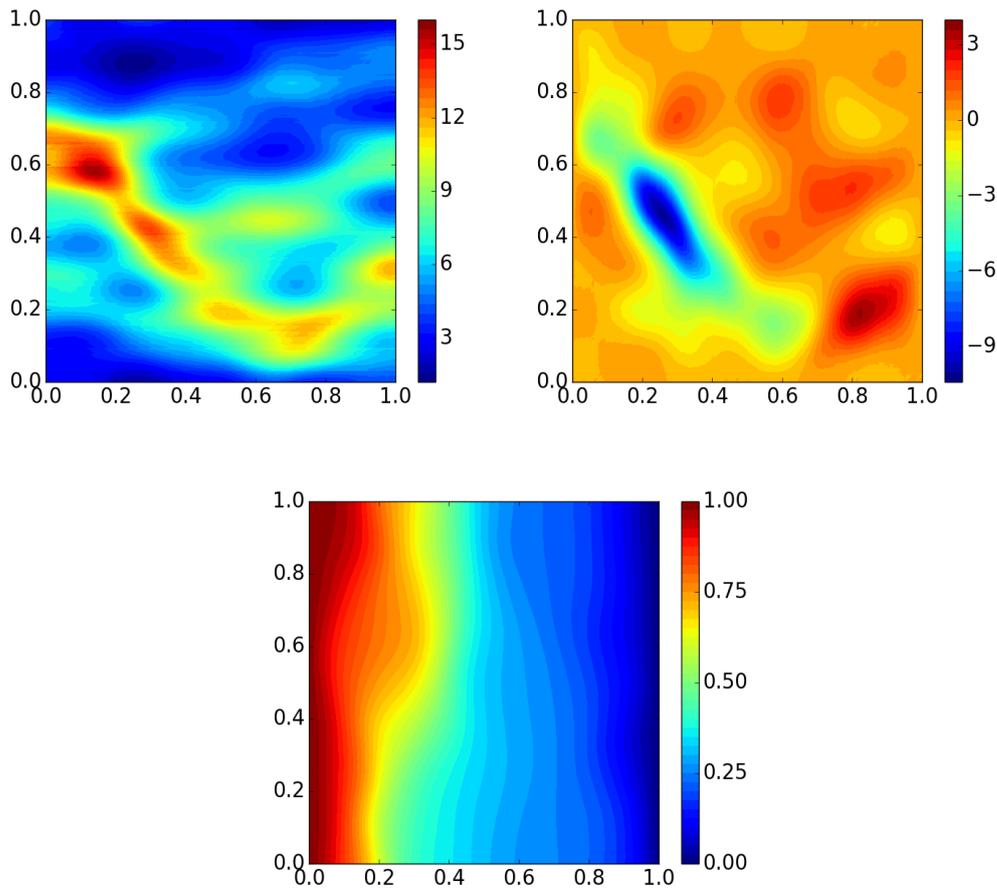


Figure 25: Functions  $\mathbb{E}[\mathbf{u}(\mathbf{x})]$  and  $\mathbb{E}[p(\mathbf{x})]$  computed with dimension-adaptive combination technique. The top row displays the components  $\mathbb{E}[u_1(\mathbf{x})]$  (left) and  $\mathbb{E}[u_2(\mathbf{x})]$  (right), at the bottom the pressure  $\mathbb{E}[p(\mathbf{x})]$  is shown.

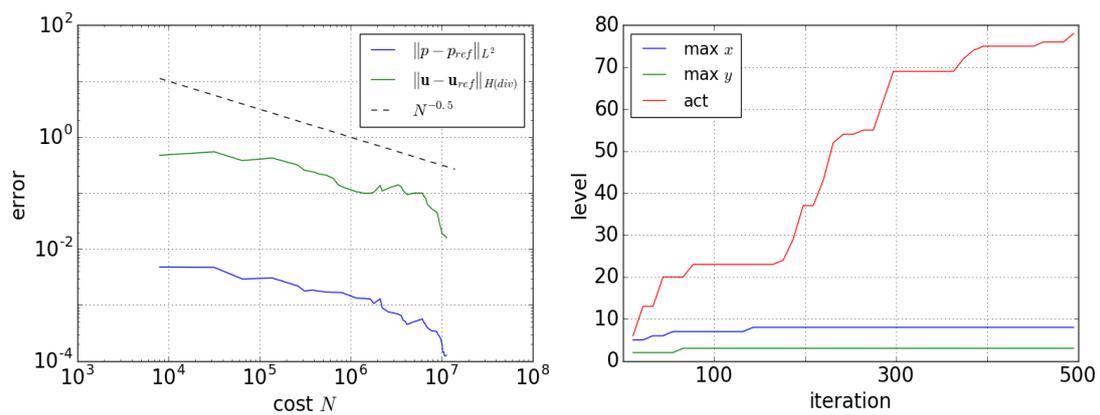


Figure 26: Convergence behavior for Darcy problem with included data points. Left: Error  $\|\mathbf{u} - \mathbf{u}_{ref}\|_{H(\text{div}, D)}$  and  $\|p - p_{ref}\|_{L^2(D)}$  with respect to the computational cost. Right: Maximal index level in  $\mathcal{I}$  after different iterations of the algorithm for spatial refinement, for all activated random variables, and the number of activated stochastic dimensions.

The algorithm proceeds for this problem similar to the other examples and the error achieves also a rate of  $\mathcal{O}(N^{-1/2})$  as illustrated in Figure 26. A difference, however, is observable in the spatial discretization. For this example, the index set  $\mathcal{I}$  includes indices corresponding to higher discretization levels of the finite element solver than in the previously. This is not surprising as the mean of the field is not constant and varies within several orders of magnitude. The variation in the field, therefore, also exists when including no stochastic parameter.

This example provides an idea of the application of the dimension-adaptive combination technique when one is interested in the modeling of the flow through porous medium for a particular application. Moreover, it shows that the algorithm also works for a non-constant mean of the random field. This example, however, portrays a very specific case, as the consideration of the field is restricted to a small subdomain and the assumption of a Gaussian covariance function is quite strong. Most likely, the usage of other covariance functions for this data set is necessary.

## 8 Summary and Outlook

In this thesis, we applied the sparse grid method in form of the combination technique in order to solve problems in uncertainty quantification which involve different numerical discretizations. We considered the computation of the first and second moments of the solution to a PDE where the diffusion coefficient is random. For the computation of these quantities of interest, three numerical approximations were looked at: a truncation of the Karhunen-Loeve expansion, which approximates the random field by a finite number of random variables, a finite element method for solving the PDE and a quadrature method for the computation of the statistics of the PDE solution, which in general is high-dimensional. To circumvent high computational costs that arise due to the curse of dimensionality, we applied the sparse grid combination technique, which combines different discretization levels of the numerical methods.

In order to find the sparse grid structure, we considered in this thesis an a-priori and a-posteriori approach. In the part dedicated to the a-priori analysis, we discussed first the regular sparse grid and later on an anisotropic version that allows combining two methods in an optimal way when the approximation power is not equal. Throughout this part, we focused separately on different aspects of the numerical approximation. We addressed how to balance the methods in a sparse grid looking at the combination of parts of the truncation, spatial and stochastic approximation.

The first focus was, after having shortly discussed the combination technique for high-dimensional integration, on a sparse grid that combines the spatial discretization with the quadrature method while keeping the number of stochastic variables fixed. We investigated the combination of the finite element solver with different quadrature rules which have different approximation power. While the regular sparse grid already provided a reduction of the computational cost, we found that it is not the optimal combination of the methods in the case when they have different approximation power. As an improved method, we applied an anisotropic version of the sparse grid, which reduces the computational cost to the cost of one method, whereby the convergence rate is preserved and is not affected by a logarithmic term.

Another aspect in the a-priori analysis concentrated on balancing the truncation, which reduces the infinite-dimensional parameter space to a finite-dimensional space, and the quadrature method applied to this set of variables, while the spatial discretization was disregarded. To this end, we looked at two methods for the quadrature: the Monte Carlo method and an anisotropic product quadrature.

The Monte Carlo method has the advantage of a dimension-independent convergence rate, but with a rather slow rate. Given the decay of importance of the stochastic parameters, we combined the Monte Carlo quadrature with varying truncation levels in a sparse grid and obtained by this a Multilevel Monte Carlo method for the integration over an infinite-dimensional parameter space that balances the methods in an optimal way with respect to the benefit-cost ratio for algebraically decaying stochastic variables.

Furthermore, we constructed a product quadrature rule based on the decay of importance of the stochastic variables, such that the convergence rate is independent of the dimension. Although the cost of the product quadrature grows heavily for approximations with high accuracies of the infinite-dimensional integral, we found in numerical examples that the product quadrature in combination with the truncation provides good approximations for moderate discretization levels and exhibit a convergence rate higher than the Monte Carlo method.

In the second part of the thesis, we determined the sparse grid structure using an a-posteriori approach. No a-priori knowledge of the regularity of the solution is required in this case, but rather a problem-dependent sparse grid structure is detected. First, we successfully adapted the algorithm for a dimension-adaptive combination technique to the problem of computing statistics of a PDE after having fixed the number of stochastic variables. In a further step, we included the truncation of the Karhunen-Loeve expansion as an additional direction in the sparse grid to the quadrature rules and PDE solver. Hence, the error made by the truncation is balanced with the quadrature error and finite element error. We constructed the algorithm such that the dimension-adaptive combination technique detects, in the course of the algorithm, the anisotropic structure within the stochastic variables and adjusts the quadrature levels to it. Numerical examples were provided to illustrate this.

This thesis concluded with the application of the dimension-adaptive combination technique to a groundwater flow problem, where typically short correlation lengths lead to high-dimensional problems. The algorithm was applied to the Darcy flow problem in mixed formulation, whereby the hydraulic conductivity was modeled with a lognormal random field. We considered fields with different smoothness for which the algorithm included the necessary stochastic variables consistent with the decay of the coefficients in the Karhunen-Loeve expansion. We restricted the application to sediment layers where the hydraulic conductivity can be modeled by a lognormal distribution, while for other formations a different model for the random field needs to be considered.

The application of the sparse grid method, either as an a-priori or a-posteriori approach, led to an improvement for the computation of the statistics of the PDE solution compared to the straightforward combination of the numerical methods. Nevertheless, different aspects seem worth investigating in future research.

A possibility for an improvement for the combination of an anisotropic quadrature with the truncation level for the integration of a function that depends on infinitely many variables might be provided by considering a sparse version. Similar to the product quadrature, the number of points for each direction would be based on the decrease of importance of the variables, but the univariate quadratures would then be combined in a sparse grid. If this allows the construction of a sequence of quadrature methods whose approximation power and cost form geometric sequences, the quadrature can be then balanced with the truncation in an optimal way similar to the Monte Carlo method.

Furthermore, as we considered only parts of the truncation, finite element, and quadrature method all at once in the a-priori analysis, an open question still is how to balance all three numerical approximations in a sparse grid such that the cost of computing the quantities of interest is reduced, but the accuracy is not much affected. While the dimension-adaptive combination technique answers this question for a given problem, the theoretical results have yet to be proven for a class of functions. The analysis in the first part of this thesis can be used to that end, but additionally a mixed regularity of the solution between the truncation, finite element and quadrature method has to be examined.

In the case where a Gaussian quadrature is applied, which potentially provides the fastest convergence rate for the quadrature, we observed that the convergence of the combination technique is limited by the convergence of the finite element solver. For a two-dimensional domain in particular, this reduces the fast convergence of the Gaussian quadrature. Hence, the Monte Carlo method, which is often applied, is usually not too restrictive.

One option to avoid a reduction of the convergence rate due to the dimension of the spatial domain is to consider a sparse grid for solving the PDE. For that, it needs to be guaranteed that the solution  $u(\mathbf{y})$  possesses mixed derivatives with respect to the spatial variables. For elliptic PDEs on a smooth enough domain, the regularity of the solution is determined by the right-hand side of the PDE. Thus, higher regularity than  $H^1(D)$  in the spatial variables for the solution is often provided. Hence, a sparse grid within the spatial variables should be possible.

Throughout this thesis, we mainly looked at numerical methods for solving the PDE which at most can achieve a convergence rate of first order. However, if the solution to the PDE is smooth enough in the spatial variable, one could consider instead a finite element method that exploits this regularity and achieves a higher convergence rate. If we know that the solution is smooth, even spectral methods could be considered, as the Gaussian quadrature converges exponentially. However, this also requires regular-shaped spatial domains, which might be the case in the study of groundwater flows, but are restrictive in other cases.

This thesis only addressed how to apply a sparse grid for the computation of the expectation and second moment of a PDE. Other quantities of interest might require additional or different approximations. For example, if the quantities of interest can not be exactly computed based on a surrogate function of  $u(\mathbf{x}, \mathbf{y})$ , an additional approximation needs to be considered. In general, the sparse grid approach is not limited to the study of our model problem. As uncertainty quantification is a broad area, sparse grid methods for other problems with several approximations could be considered where several numerical methods are combined in order to reduce the cost.

## References

- [1] Ralph C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*. Computational science and engineering series. Society for Industrial and Applied Mathematics, 2013.
- [2] T.J. Sullivan. *Introduction to Uncertainty Quantification*. Springer International Publishing, 2015.
- [3] Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization. *SIAM Review*, 60(3):550–591, 2018.
- [4] Gabriel J. Lord, Catherine E. Powell, and Tony Shardlow. *An Introduction to Computational Stochastic PDEs*. Number 50 in Cambridge texts in applied mathematics. Cambridge University Press, 2014.
- [5] Michael Griebel and Helmut Harbrecht. On the Construction of Sparse Tensor Product Spaces. *Mathematics of Computation*, 82(282):975–994, 2013.
- [6] Michael Griebel, Helmut Harbrecht, and Michael D. Multerer. Multilevel Quadrature for Elliptic Parametric Partial Differential Equations in Case of Polygonal Approximations of Curved Domains. *SIAM Journal on Numerical Analysis*, 58(1):684–705, 2020.
- [7] T. Gerstner and M. Griebel. Dimension Adaptive Tensor Product Quadrature. *Computing*, 71(1):65–87, 2003.
- [8] A. Rüttgers and M. Griebel. Multiscale Simulation of Polymeric Fluids Using the Sparse Grid Combination Technique. *Applied Mathematics and Computation*, 319:425–443, 2018.
- [9] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
- [10] Philip J Davis, Philip Rabinowitz, and Werner Rheinbolt. *Methods of Numerical Integration*. Elsevier Science, 2014.
- [11] Martin Hanke-Bourgeois. *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Vieweg+Teubner, 2009.
- [12] Andrea Barth, Christoph Schwab, and Nathaniel Zollinger. Multi-level Monte Carlo Finite Element Method for Elliptic PDEs with Stochastic Coefficients. *Numerische Mathematik*, 119(1):123–161, 2011.
- [13] Art B. Owen. *Monte Carlo Theory, Methods and Examples*. 2013. <https://statweb.stanford.edu/~owen/mc/>.
- [14] Harald Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. Number 63 in CBMS-NSF regional conference series in applied mathematics. Society for Industrial and Applied Mathematics, 1992.

- [15] Markus Siebenmorgen. *Quadrature Methods for Elliptic PDEs with Random Diffusion*. PhD thesis, Basel, 2015.
- [16] Jörg Waldvogel. Fast Construction of the Fejér and Clenshaw–Curtis Quadrature Rules. *BIT Numerical Mathematics*, 46(1):195–202, 2006.
- [17] Giovanni Monegato. An overview of the computational aspects of Kronrod quadrature rules. *Numerical Algorithms*, 26(2):173–196, 2001. Publisher: Springer US, New York, NY.
- [18] Thomas N. L. Patterson. The Optimum Addition of Points to Quadrature Formulae. *Mathematics of Computation*, 22(104):847–s31, 1968.
- [19] Ivo Babuška, Fabio Nobile, and Raúl Tempone. A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.
- [20] Marcel Bieri. A Sparse Composite Collocation Finite Element Method for Elliptic SPDEs. *SIAM Journal on Numerical Analysis*, 49(6):2277–2301, 2011.
- [21] Abdul-Lateef Haji-Ali, Helmut Harbrecht, Michael Peters, and Markus Siebenmorgen. Novel Results for the Anisotropic Sparse Grid Quadrature. *Journal of Complexity*, 47:62–85, 2018.
- [22] Shuhuang Xiang and Folkmar Bornemann. On the Convergence Rates of Gauss and Clenshaw-Curtis Quadrature for Functions of Limited Regularity. *SIAM Journal on Numerical Analysis*, 50(5):2581–2587, 2012.
- [23] Helmut Harbrecht, Michael Peters, and Markus Siebenmorgen. Multilevel Accelerated Quadrature for PDEs with Log-Normally Distributed Diffusion Coefficient. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):520–551, 2016.
- [24] Richard E Bellman. *Adaptive Control Processes*. Princeton University Press, 1961.
- [25] Dietrich Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Elasticity Theory*. Cambridge University Press, 3rd ed edition, 2007.
- [26] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer New York, 2008.
- [27] Alfio Quarteroni. *Numerical Models for Differential Problems*. Springer Milan, 2014.
- [28] FEniCS Project. <https://fenicsproject.org/>.
- [29] Markus Bachmayr, Albert Cohen, Ronald DeVore, and Giovanni Migliorati. Sparse Polynomial Approximation of Parametric Elliptic PDEs. Part II: Lognormal Coefficients. *ESAIM: Mathematical Modelling and Numerical Analysis*, 51(1):341–363, 2017.

- [30] Frances Y. Kuo, Christoph Schwab, and Ian H. Sloan. Quasi-Monte Carlo Finite Element Methods for a Class of Elliptic Partial Differential Equations with Random Coefficients. *SIAM Journal on Numerical Analysis*, 50(6):3351–3374, 2012.
- [31] Michael Griebel and Guanglian Li. On the Decay Rate of the Singular Values of Bivariate Functions. *SIAM Journal on Numerical Analysis*, 56(2):974–993, 2018.
- [32] Christoph Schwab and Radu Alexandru Todor. Karhunen–Loève Approximation of Random Fields by Generalized Fast Multipole Methods. *Journal of Computational Physics*, 217(1):100–122, 2006.
- [33] I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, Ch. Schwab, and I. H. Sloan. Quasi-Monte Carlo Finite Element Methods for Elliptic PDEs with Lognormal Random Coefficients. *Numerische Mathematik*, 131(2):329–368, 2015.
- [34] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive computation and machine learning. MIT Press, 2006.
- [35] Andreas Keese. *Numerical Solution of Systems with Stochastic Uncertainties: A General Purpose Framework for Stochastic Finite Elements*. PhD thesis, Braunschweig, 2004.
- [36] Smolyak, S. A. Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions. *Doklady Akademii Nauk*, 1963.
- [37] Christoph Zenger. Sparse Grids. In *Parallel Algorithms for Partial Differential Equations*, number v. 31 in Notes on numerical fluid mechanics, pages 241–251. Vieweg, 1991.
- [38] Hans-Joachim Bungartz and Michael Griebel. Sparse Grids. *Acta Numerica*, 13:147–269, 2004.
- [39] F. Nobile, R. Tempone, and C. G. Webster. An Anisotropic Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.
- [40] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger. A Multi-level Stochastic Collocation Method for Partial Differential Equations with Random Input Data. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):1046–1074, 2015.
- [41] Abdul-Lateef Haji-Ali, Fabio Nobile, Lorenzo Tamellini, and Raul Tempone. Multi-Index Stochastic Collocation for random PDEs. *Computer Methods in Applied Mechanics and Engineering*, 306:95–122, 2016.

- [42] Oliver G. Ernst, Björn Sprungk, and Lorenzo Tamellini. Convergence of Sparse Collocation for Functions of Countably Many Gaussian Random Variables (with Application to Elliptic PDEs). *SIAM Journal on Numerical Analysis*, 56(2):877–905, 2018.
- [43] Frances Y. Kuo, Robert Scheichl, Christoph Schwab, Ian H. Sloan, and Elisabeth Ullmann. Multilevel quasi-Monte Carlo methods for lognormal diffusion problems. *Mathematics of Computation*, 86(308):2827–2860, 2017.
- [44] Frances Y. Kuo, Christoph Schwab, and Ian H. Sloan. Multi-level Quasi-Monte Carlo Finite Element Methods for a Class of Elliptic PDEs with Random Coefficients. *Foundations of Computational Mathematics*, 15(2):411–449, 2015.
- [45] Josef Dick, Frances Y. Kuo, and Ian H. Sloan. High-dimensional Integration: The Quasi-Monte Carlo Way. *Acta Numerica*, 22:133–288, 2013.
- [46] Fred J. Hickernell, Thomas Müller-Gronbach, Ben Niu, and Klaus Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on RN. *Journal of Complexity*, 26(3):229–254, 2010.
- [47] Jochen Garcke. Sparse Grids in a Nutshell. In Jochen Garcke and Michael Griebel, editors, *Sparse Grids and Applications*, volume 88 of *Lecture Notes in Computational Science and Engineering*, pages 57–80. Springer Berlin Heidelberg, 2012.
- [48] Erich Novak and Klaus Ritter. The Curse of Dimension and a Universal Method For Numerical Integration. In Günther Nürnberger, Jochen W. Schmidt, and Guido Walz, editors, *Multivariate Approximation and Splines*, pages 177–187. Birkhäuser Basel, 1997.
- [49] Thomas Gerstner and Michael Griebel. Numerical Integration Using Sparse Grids. *Numerical Algorithms*, 18(3/4):209–232, 1998.
- [50] Michael Griebel, Michael Schneider, and Christoph Zenger. A Combination Technique for the Solution of Sparse Grid Problems. In P. de Groen and R. Beauwens, editors, *Iterative Methods in Linear Algebra*, pages 263–281. IMACS, Elsevier, 1992.
- [51] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients. *Computing and Visualization in Science*, 14(1):3–15, 2011.
- [52] Michael B. Giles. Multilevel Monte Carlo Path Simulation. *Operations Research*, 56(3):607–617, 2008.
- [53] Joakim Beck, Fabio Nobile, Lorenzo Tamellini, and Raúl Tempone. Convergence of Quasi-optimal Stochastic Galerkin Methods for a Class of PDEs with Random Coefficients. *Computers & Mathematics with Applications*, 67(4):732–751, 2014.

- [54] Alexandros Gilch, Michael Griebel, and Jens Oettershagen. Multilevel Approximation for Generalized Method of Moments Estimators. Submitted to International Journal for Uncertainty Quantification. Available as INS Preprint No. 2006, 2020.
- [55] Brian R. Hunt, Tim Sauer, and James A. Yorke. Prevalence: A Translation-invariant “almost every” on Infinite-dimensional Spaces. *Bulletin of the American Mathematical Society. New Series*, 27(2):217–238, 1992.
- [56] Nelson Dunford. *Linear Operators*. New York, 1958.
- [57] Albert Cohen, Ronald DeVore, and Christoph Schwab. Convergence Rates of Best N-term Galerkin Approximations for a Class of Elliptic sPDEs. *Foundations of Computational Mathematics*, 10(6):615–646, 2010.
- [58] Albert Cohen, Ronald Devore, and Christoph Schwab. Analytic Regularity and Polynomial Approximation of Parametric and Stochastic Elliptic PDE’s. *Analysis and Applications*, 09(01):11–47, 2011.
- [59] Alexander Rüttgers. *Multiscale Simulation of Polymeric Fluids using Sparse Grids*. PhD thesis, Bonn, 2016.
- [60] F. Nobile, L. Tamellini, and R. Tempone. Convergence of Quasi-optimal Sparse-grid Approximation of Hilbert-space-valued Functions: Application to Random Elliptic PDEs. *Numerische Mathematik*, 134(2):343–388, 2016.
- [61] Fabio Nobile, Lorenzo Tamellini, Francesco Tesei, and Raúl Tempone. An Adaptive Sparse Grid Algorithm for Elliptic PDEs with Lognormal Diffusion Coefficient. In Jochen Garcke and Dirk Pflüger, editors, *Sparse Grids and Applications - Stuttgart 2014*, volume 109, pages 191–220. Springer International Publishing, 2016.
- [62] Alexander D. Gilbert, Frances Y. Kuo, Dirk Nuyens, and Grzegorz W. Wasilkowski. Efficient implementations of the Multivariate Decomposition Method for approximating infinite-variate integrals. *arXiv:1712.06782 [cs, math]*, 2018.
- [63] Jacob Bear and Alexander H.-D. Cheng. *Modeling Groundwater Flow and Contaminant Transport*. Springer Netherlands, 2010.
- [64] SPE Comparative Solution Project - Description of Model 2. <https://www.spe.org/web/csp/datasets/set02.htm>.
- [65] Huangxin Chen, Amgad Salama, and Shuyu Sun. Adaptive Mixed Finite Element Methods for Darcy Flow in Fractured Porous Media. *Water Resources Research*, 52(10):7851–7868, 2016.
- [66] Knut-Andreas Lie. *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST)*. Cambridge University Press, 1 edition, 2019.

- [67] Sarah Osborn, Panayot S. Vassilevski, and Umberto Villa. A Multilevel, Hierarchical Sampling Technique for Spatially Correlated Random Fields. *SIAM Journal on Scientific Computing*, 39(5):s543–s562, 2017.
- [68] I. G. Graham, R. Scheichl, and E. Ullmann. Mixed Finite Element Analysis of Lognormal Diffusion and Multilevel Monte Carlo Methods. *Stochastics and Partial Differential Equations Analysis and Computations*, 4(1):41–75, 2016.
- [69] Joakim Beck, Fabio Nobile, Lorenzo Tamellini, and Raúl Tempone. A Quasi-optimal Sparse Grids Procedure for Groundwater Flows. In Mejdí Azañez, Henda El Fekih, and Jan S. Hesthaven, editors, *Spectral and High Order Methods for Partial Differential Equations - ICOSAHOM 2012*, volume 95, pages 1–16. Springer International Publishing, 2014.
- [70] Luca Bergamaschi, Stefano Mantica, and Fausto Saleri. Mixed Finite Element Approximation of Darcy’s Law In Porous Media. *TR CRS4-AppMath-94-17*, 1994.
- [71] Paul Constantine. Random Field Simulation.  
<https://de.mathworks.com/matlabcentral/fileexchange/27613-random-field-simulation>, downloaded October 2020.