Doctoral Thesis

Sparse perturbation algorithms for elliptic PDE's with stochastic data

Author(s):
Todor, Radu Alexandru

Publication Date:
2005

Permanent Link:
https://doi.org/10.3929/ethz-a-005151392

Rights / License:
In Copyright - Non-Commercial Use Permitted
Sparse Perturbation Algorithms
for Elliptic PDE's with Stochastic Data

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of
Doctor of Mathematics

presented by
RADU ALEXANDRU TODOR
Dipl. Math. University of Bucharest
born November 12, 1973
citizen of Romania

accepted on the recommendation of
Prof. Dr. Christoph Schwab, examiner
Prof. Dr. Ralf Hiptmair, co-examiner
Prof. Dr. Reinhold Schneider, co-examiner

2005
Dank

Die vorliegende Arbeit ist unter der Leitung von Herrn Prof. Christoph Schwab entstanden. Ich danke ihm ganz herzlich für seine stete Unterstützung, seine hilfreichen und motivierenden Ratschläge, und seine immer sehr engagierte Betreuung meiner Doktorarbeit.

Herrn Prof. Ralf Hiptmair und Herrn Prof. Reinhold Schneider danke ich für ihre konstruktiven Bemerkungen und die Übernahme des Koreferats.

Ferner gilt mein Dank meiner Familie, die mich immer liebevoll unterstützt hat, sowie meinen Kolleginnen und Kollegen aus dem Seminar für Angewandte Mathematik, die mir den Aufenthalt an der ETH angenehm gemacht haben.

Financial support by the European Union (contract number HPRN-CT-2002-00286) and the Swiss Federal Office for Science and Education (grant BBW 02.0418), within the IHP network 'Breaking Complexity', is gratefully acknowledged.
Seite Leer /
Blank leaf
Abstract

In this work we investigate elliptic partial differential equations with stochastic data. We develop fully deterministic algorithms using statistics of the data (expectation and/or higher order moments) as input, for the computation of similar statistics of the stochastic solution. We introduce a sparsified higher order moment representation which, in the context of a classical perturbation method enables us to formulate efficient algorithms for the computation of solution moments, using sparse grids and wavelet techniques. Assuming analytic spatial regularity of the data fluctuation, we prove almost linear complexity for these algorithms.

Choosing stationary diffusion in a bounded domain as a model problem, we discuss in Chapter 1 the setting, basic notations and definitions, plus the main results.

We devote Chapter 2 to the case of a stochastic source term and deterministic diffusion coefficient. We derive deterministic moment equations and use sparse grids discretization to preserve almost optimal convergence rates. We propose a new solution algorithm for the resulting linear system given by a well-conditioned (due to the use of a wavelet basis), yet fully populated matrix with a tensor product block structure, to achieve log-linear complexity (in the number of degrees of freedom) despite the high dimensionality of the moment problem.

In Chapter 3 we address the more general case of a stochastic diffusion coefficient and stochastic source term. We show that the expectation of the stochastic solution can be computed starting from the higher order moments of the data. We introduce approximate sparsified representations of these moments and show that using them as input in the computation of the solution expectation results in an algorithm of almost linear complexity, under the analyticity assumption mentioned above. We finally combine this algorithm with the one developed in Chapter 2 to obtain similar, efficient algorithms for the computation of higher order moments of the stochastic solution.

We conclude by discussing in Chapter 4 possible further theoretical developments and extensions of the results presented in this thesis.
Kurzfassung


In Kapitel 1 wird das Modelproblems (stationäre Diffusion im beschränkten physikalischen Gebiet), die grundlegenden Konzepte und Definitionen eingeführt, sowie die Hauptresultate erläutert.

Der Fall einer stochastischen Quelle mit deterministischem Diffusionskoeffizient wird in Kapitel 2 behandelt. Hoch dimensionale, deterministische Momentgleichungen werden hergeleitet, mit Hilfe von dünnen Gittern und Waveletbasen diskretisiert, was zu gut konditionierten, allerdings voll besetzten linearen Systemen führt. Ein Lösungsverfahren wird entwickelt, das die Tensorproduktstruktur des ursprünglichen Momentproblems ausnutzt, um log-lineare Gesamtkomplexität zu erreichen.

Für das vollständig stochastische Diffusionsproblem (sowohl Quelle als auch Diffusionskoeffizient werden als Zufallsfelder modelliert) wird in Kapitel 3 ein Algorithmus zur Berechnung des Erwartungswertes entwickelt. Höhere Momente der stochastischen Daten werden algorithmisch in eine Approximation des Erwartungswertes der stochastischen Lösung umgesetzt. Im Falle einer im Ort regulären Fluktuation der Daten und unter Verwendung der effizienten Darstellungen der Momente als Eingabe wird erneut fast lineare Gesamtkomplexität bewiesen. Ferner wird dieser Algorithmus mit dem in Kapitel 2 beschriebenen Verfahren kombiniert, was auch eine effiziente Berechnung der höheren Momente ermöglicht.

Im letzten Kapitel werden Möglichkeiten zur Verallgemeinerung der zuvor entwickelten Theorie aufgezeigt.
Seite Leer / Blank leaf
# Contents

## Dank

Abstract

Kurzfassung

1 Introduction
   1.1 Problem Formulation ........................................... 1
   1.2 Main Results .................................................. 4
   1.3 Notations and Definitions .................................... 8

2 Stochastic Source Term
   2.1 Generalities .................................................... 11
   2.2 The Moment Equation .......................................... 13
   2.3 Discretization with Sparse FEM ................................ 14
   2.4 Iterative Solution and Complexity ............................ 17

3 Stochastic Coefficient
   3.1 Random Fields .................................................. 27
   3.2 Discretization .................................................. 30
   3.3 Neumann Series Decomposition ................................ 31
   3.4 Discrete Mixed Moment Problem ................................. 33
      3.4.1 Mixed Moment Algorithm .................................. 36
      3.4.2 Sparsification ............................................. 41
      3.4.3 Complexity ............................................... 43
   3.5 Mean Field Computation ....................................... 44
   3.6 Computation of Higher Order Moments ........................ 45
      3.6.1 Discretization ............................................. 46
      3.6.2 Algorithms ............................................... 48

4 Conclusions and Open Issues ..................................... 57
### 5 Appendix

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 A Trace Lemma</td>
<td>59</td>
</tr>
<tr>
<td>5.2 A Fast Decaying Sequence</td>
<td>61</td>
</tr>
</tbody>
</table>

### Bibliography

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliography</td>
<td>63</td>
</tr>
</tbody>
</table>

### Curriculum Vitae

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curriculum Vitae</td>
<td>67</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Problem Formulation

The rapid development of (adaptive) algorithms, hardware and software in recent years has made the accurate numerical solution of elliptic partial differential equations a routine matter in many (but not all) engineering applications. Despite this, results of accurate finite element (FE) computations often deviate significantly from the responses of the physical system under consideration. Having eliminated the discretization error, and assuming that the modelling error inherent in the selected partial differential equations (pde's) is negligible (i.e. that the adopted pde's precisely describe the physics of the system under consideration), the gap between simulation and observation must be due to uncertainty in the input data.

One task for applied mathematics and scientific computing is therefore (see e.g. [GS91], [ER03]) to develop tools for uncertainty processing, i.e. the systematic and quantitative numerical representation of uncertainty in input data and its propagation to the output of a finite element simulation. This requires new developments in several areas of applied mathematics and engineering: input parameters are replaced by random variables resp. by random fields with known or estimated statistics, and traditional deterministic finite element methods are reformulated to allow for randomness in input data and solution.

In this work, we focus on this latter aspect, i.e. the formulation, design and analysis of deterministic finite element solutions of elliptic pde's with stochastic data.

We consider as a model problem diffusion in a bounded domain $D \subset \mathbb{R}^d$, where the data (source term $f$ and diffusion coefficient $a$) and therefore the solution $u$ too, are allowed to depend randomly on a new parameter $\omega \in \Omega$,

$$-\text{div}(a(\cdot, \omega) \nabla u(\cdot, \omega)) = f(\cdot, \omega) \quad \text{in } H^{-1}(D) \quad P\text{-a.e. } \omega \in \Omega. \quad (1.1)$$

Here we assume the sample set $\Omega$ to be equipped with a probability measure $P$ quantifying the uncertainty and giving the probability of occurrence for each event $\sigma$ in a set $\Sigma \subset 2^\Omega$ of admissible events (that is, $(\Omega, \Sigma, P)$ probability space).
Identifying the data \(a, f\) as well as the solution \(u\) with random vector fields we rewrite (1.1) as

\[
\Delta_{a(\omega)}u(\omega) = f(\omega) \quad \text{in } H^{-1}(D) \quad P\text{-a.e. } \omega \in \Omega. \tag{1.2}
\]

Both (1.1) and (1.2) are to be understood here and throughout this work variationally in \(H^1_0(D)\), that is,

\[
\int_D a(x, \omega) \nabla u(x) \cdot \nabla v(x) \, dx = \langle f(\cdot, \omega), v(\cdot) \rangle_{H^{-1}(D), H^1_0(D)} \quad \forall v \in H^1_0(D) \quad P\text{-a.e. } \omega \in \Omega.
\]

Under the uniform ellipticity assumption (\(\lambda\) denotes here the Lebesgue measure)

\[
a_- \leq a(x, \omega) \leq a_+ \quad \lambda \times P\text{-a.e. } (x, \omega) \in D \times \Omega,
\]

(1.1) is well-posed in \(L^0(\Omega, H^1_0(D))\) for data \(f \in L^0(\Omega, H^{-1}(D))\) and \(a \in L^0(D \times \Omega)\) (here \(L^0\) denotes measurability). For equation (1.1) we then formulate the following

**Problem.** Given statistical information w.r.t. \(\omega\) about the data \((a, f)\), compute statistical information w.r.t. \(\omega\) about \(u\) solution to (1.1) (statistical information comprises first and higher order moments, variance, probability distribution function etc.).

Since for a problem with stochastic data knowing all joint probability densities (i.e. complete statistical information) is in practice hardly the case, reasonable assumptions can be made only on some data statistics. Defining for any \(k \in \mathbb{N}_+\) the moment of order \(k\) of a measurable \(u : D \times \Omega \to \mathbb{R}\) by

\[
\mathcal{M}^k(u) : D^k := D \times D \times \cdots \times D \to \mathbb{R}
\]

\[
\mathcal{M}^k(u)(x_1, x_2, \ldots, x_k) := \int_\Omega u(x_1, \omega)u(x_2, \omega)\cdots u(x_k, \omega) \, dP(\omega)
\]

whenever the integral exists, the moment version of the problem above reads,

**Moment Problem.** Given moments w.r.t. \(\omega\) of the data \((a, f)\), compute moments w.r.t. \(\omega\) of \(u\) solution to (1.1) (most relevant are first and second order moments).

Several approaches have been proposed in the literature to treat these two problems numerically, of which we mention and briefly discuss here the Monte Carlo (MC), the perturbation and the stochastic Galerkin (SG) methods.

The MC simulation is probably the simplest and most natural approach, based on sampling the coefficients and solving deterministic problems by a standard (e.g. Galerkin) finite element method (see, e.g. [BTZ04]). The computation is fully parallelizable (therefore easy to code), but requires knowledge of data distribution everywhere in the physical domain,
1.1. PROBLEM FORMULATION

a very strong assumption. Besides, the method is in general considered to converge slowly due to the statistical error generated by sampling, although the overall complexity can be easily shown to behave *polynomially* in inverse accuracy (see Definition 1.1).

The perturbation approach is widely used in engineering applications (see [KH92] and references therein). There are several versions, of which we mention here only the *Neumann expansion* (see [Kel64]), the *first order second moment (FOSM)* (see [DW81]), and the *Wiener/generalized polynomial chaos (W/gPC)* (see [Wie38], [Sch00], [XK02]) methods. In one of its mathematically rigorous forms the perturbation method consists in inverting the operator with random coefficients as a Neumann series around its deterministic mean and solving then sequentially a large number of deterministic problems with different source terms. Its good performance has been demonstrated in practice (at least for small fluctuations), but only *superalgebraic complexity* estimates have been obtained in theory (see e.g. [BC02]). In this work we also adopt the Neumann perturbation approach and, after carrying out a careful error analysis, we propose an improved version, involving also *sparse grids and wavelet* techniques, and leading, under additional but reasonable *regularity assumption on the fluctuation*, to moment computation algorithms of *subalgebraic, nearly optimal complexity* (see Definition 1.1 below). Note that the fluctuation regularity assumption allows the control of the data truncation error and acts here therefore as a substitute for the standard *moment closure hypothesis*.

The parametrization of uncertainty is one of the key points in the numerical treatment of problems with stochastic data. The number of parameters used to describe the random fields has in general a huge impact on the algorithm complexity. A random field expansion of Karhunen-Löeve type separating the deterministic and stochastic variables at an optimal approximation rate (see e.g. [Loè77], [Loè78]) has become in recent years a standard procedure to transform the original stochastic problem into a parametric deterministic one, where the parameter involved belongs to a high dimensional space.

The SG method interprets then the transformed problem variationally in both the parameter and the physical variable and uses a standard (Galerkin) FEM to produce a numerical solution (note the need for numerical integration schemes in high dimensional domains). Backward substitution gives the numerical solution to the original stochastic problem and postprocessing is required to obtain the desired statistical information. Just as in the case of a MC simulation, very detailed information on the data distribution is needed, and the finite elements in use today to treat the transformed problem can only ensure *superalgebraic complexity* rates (see e.g. [BTZ04], [FST05], [MK05]).

An interesting and promising alternative to the SG method would be to essentially apply the same complexity reduction method used here (see section 3.4.2) in the context of reconstruction techniques (collocation, interpolation) for the solution of the transformed
problem, by taking into account its analytic parameter dependence. This leads to results comparable with those presented in this work - \textit{nearly optimal complexity} under qualitatively similar assumptions.

**Definition 1.1.** We say that an algorithm for the computation of $M^k(u)$ with $u$ solution to \eqref{eq:1.1} has \textit{nearly optimal complexity} if, assuming that the deterministic diffusion problem in $D$ with data given by the expectations of $a$ and $f$ can be solved with an accuracy $\varepsilon$ using $O(\varepsilon^{-\beta})$ operations and memory (for some $\beta > 0$, asymptotically as $\varepsilon \to 0$), the algorithm requires at most $O(\varepsilon^{-\beta-o(1)})$ operations and memory for an $\varepsilon$-accurate computation of $M^k(u)$.

In other words, an algorithm for the computation of $M^k(u)$ (where $u$ solves a stochastic model problem) is nearly optimal if the computational effort required is equivalent, asymptotically as the prescribed accuracy $\varepsilon$ goes to 0, to the effort needed to approximately solve one deterministic model problem. For example, if the deterministic model problem can be solved in linear complexity ($\beta = 1$), e.g. by a multigrid method, then a nearly optimal algorithm for $M^k(u)$ has (nearly) linear complexity, too.

**Example 1.2.** An algorithm for the computation of $M^1(u)$ with $u$ solution to \eqref{eq:1.1}, which amounts to solving $\varepsilon^{-o(1)}$ deterministic diffusion problems in $D$ for an accuracy of $\varepsilon \to 0$ and with data as regular as the expectations of $a$, $f$, has nearly optimal complexity.

We mention that the mean field computation algorithm we are going to present is of the type mentioned in the example above and will be developed around some very classical ideas of the perturbation approach. Besides, it is almost fully parallelizable and its implementation can be based on any preexistent FE solver of the deterministic model (diffusion) problem in $D$ (to be called with many different r.h.s.'s).

### 1.2 Main Results

We give in the following a summary of the main results contained in this work. We emphasise that the main novelty consists in improved perturbation algorithms of \textit{nearly optimal complexity} (thus more efficient than MC, standard perturbation or SG methods, under similar assumptions).

The case of a \textit{deterministic} diffusion coefficient $a$ in \eqref{eq:1.1} (that is, depending only on the \textit{physical} variable $x$, $a(x, \omega) = a(x) \forall x \in D$) with a \textit{stochastic} source term $f$ (that is, depending on both the physical and the \textit{stochastic} variables $x$ and $\omega$ respectively) is a key step in the development of solution algorithms for \eqref{eq:1.1} using a perturbation approach.
1.2. MAIN RESULTS

This analysis has been carried out in [ST03a], [ST03b]. We briefly review next the main results obtained in [ST03b].

Defining the solution operator to the model problem, $\Delta_a := -\text{div}(a\nabla) : H^0_0(D) \to H^{-1}(D)$, and

$$
\Delta_a \otimes \Delta_a \otimes \cdots \otimes \Delta_a : H^0_0(D^k) \to H^{-1}(D^k)
$$

where

$$
H^s(D^k) := H^s(D) \otimes H^s(D) \otimes \cdots \otimes H^s(D) \quad \forall s \in [-1, \infty[,
$$

$$
H^0_0(D^k) := H^0_0(D) \otimes H^0_0(D) \otimes \cdots \otimes H^0_0(D)
$$

are the anisotropic Sobolev spaces of mixed highest derivatives (e.g. $H^1(D^2) \simeq \{v(x,y) \in L^2(D \times D) | \nabla_x v, \nabla_y v, \nabla_x \nabla_y v \in L^2\}$), a deterministic equation for $\mathcal{M}^k(u)$ in terms of $\mathcal{M}^k(f)$ is

$$(\Delta_a \otimes \Delta_a \otimes \cdots \otimes \Delta_a)\mathcal{M}^k(u) = \mathcal{M}^k(f) \quad \text{in } H^{-1}(D^k).$$

This deterministic moment problem for $\mathcal{M}^k(u)$ in terms of $\mathcal{M}^k(f)$ shows that we can trade randomness in the original pde at the expense of high dimension in the problem domain. Employing a sparse tensor product discretization and a wavelet preconditioning procedure, the corresponding discrete moment problem is shown to lead to a well-conditioned linear system. Further, a solution algorithm for the resulting system (given in general by a fully populated matrix, but having a special blockwise tensor product structure) is developed to ensure log-linear (in $N$, number of dof's in $D$) number of operations and memory requirements. The high dimensional $k$-th moment problem is therefore solvable numerically in work essentially proportional to that of the mean field equation, i.e. for $k = 1$, up to logarithmic terms. In other words, nearly optimal complexity is achieved by the algorithm presented in [ST03b] in the case of a deterministic diffusion coefficient.

In chapter 3 we take up the case of a stochastic coefficient. The main difficulty is now that, in contrast to the case of a deterministic coefficient, there is no obvious pde with $\mathcal{M}^k(u)$ as a solution. We start with the simplest such problem, that of computing $\mathcal{M}^1(u)$, the mean field (or expectation) of $u$ solution to (1.1). The idea is, as in the widely used perturbation approach, to use a representation (decomposition)

$$
a(x, \omega) = e(x) + r(x, \omega) = \text{'deterministic expectation' + 'random fluctuation'}
$$

where the fluctuation is assumed to be smaller than the expectation in the sense that

$$
0 \leq \sup_{x \in D} \frac{\|r(x, \cdot)\|_{L^\infty(D)}}{e(x)} < 1.
$$

Choosing $e$ appropriately, it is easily seen that this condition is (almost) equivalent to the ellipticity of $\Delta_a$. 
We further assume the fluctuation $r$ to be analytic in the physical domain $D \subset [-1, 1]^d$, but emphasize that this condition can be relaxed to only piecewise finite Sobolev regularity,

$$r \in A([-1, 1]^d, L^\infty(\Omega)).$$

The analyticity assumption, which is satisfied e.g. if the covariance of $a$ is Gaussian, ensures the existence of a fluctuation expansion as a fast (quasi-exponential) convergent series separating the deterministic and stochastic variables,

$$r(x, \omega) = \sum_{m \in \mathbb{N}_+} \phi_m(x)X_m(\omega), \quad \|\phi_m \otimes X_m\|_{L^\infty(D \times \Omega)} \lesssim e^{-O(m^{1/d})} \quad \forall m \in \mathbb{N}_+. \quad (1.3)$$

Such a representation is clearly not unique - one can choose e.g. the Karhunen-Loève (KL) expansion of $r$, or $(\phi_m)_{m \in \mathbb{N}_+}$ to be the Legendre polynomials in $[-1, 1]^d$. In both cases the deterministic part $(\phi_m)_{m \in \mathbb{N}_+}$ of the expansion (1.3) is available (by an additional eigenvalue computation for the integral operator associated to the kernel $M^2(r)$, in the KL case) and statistical information on the stochastic part $(X_m)_{m \in \mathbb{N}_+}$ follows by testing $r$ against $(\phi_m)_{m \in \mathbb{N}_+}$.

$$X_m(\omega) = \int_\Omega r(x, \omega)\phi_m(x)\, dP(\omega) \quad \forall m \in \mathbb{N}_+.$$  

Choosing a truncation order $M$ for the fluctuation series,

$$r_M := \sum_{m=1}^M \phi_m(x)X_m(\omega)$$

and a FE discretization level $L$ (corresponding to the FE space $V_L$ of dimension $N_L$), we define recursively, with $\Delta e = -\text{div}(e\nabla)$, the sequence $(v_{j,L})_{j \in \mathbb{N}}$ by

$$\Delta e v_{0,L}(\cdot, \omega) = f(\cdot, \omega) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega.$$  

$$\Delta e v_{j,L}(\cdot, \omega) = \text{div}(r_M(\cdot, \omega)v_{j-1,L}(\cdot, \omega)) \quad \forall j \in \mathbb{N}_+$$

The Neumann series $\sum_{j \in \mathbb{N}} v_{j,L}$ is then easily shown to converge in $H^1_0(D)$, uniformly in $\omega$, like a geometric series. Truncating it a level $n \in \mathbb{N}$ and assuming the data $e, f$ to be sufficiently regular in $D$ as to ensure a FE convergence rate $\Phi(N_L)$ uniformly in $\omega \in \Omega$, we have

$$\|M^1(u) - \sum_{j=0}^n M^1(v_{j,L})\|_{H^1_0(D)} \lesssim e^{-O(M^{1/d})} + \Phi(N_L) + e^{-O(n)}. \quad (1.4)$$

In order to balance the contributions of the three discretization steps taken so far (truncation of fluctuation series, FE in $D$, truncation of Neumann series) and to achieve an accuracy of order $\varepsilon > 0$ we choose

$$M \sim |\log \varepsilon|^d, \quad L \text{ s.t. } \Phi(N_L) \leq \varepsilon, \quad n \sim |\log \varepsilon|. \quad (1.5)$$
1.2. MAIN RESULTS

However, the computation of $\mathcal{M}^1(v_{j,L})$ for $0 \leq j \leq n$ in (1.4) is not an obvious task. We show that $\mathcal{M}^1(v_{j,L})$ can be computed exactly using any deterministic solver in $D$ at FE discretization level $L$, starting from the mixed moment

$$\mathcal{M}^{(j,1)}(r_M, f)(x, x) := \int_D r_M(x, \omega) f(x, x) \, dP(\omega) \quad x \in D^j, x \in D$$

via the following composition of $j + 1$ bounded operators defined in terms of the div, $\nabla$ operators in $D$, the trace operator on the diagonal set in $D \times D$ and the FE solution operator $\Delta_{e,L}^{-1}$ of the deterministic diffusion problem in $D$ with coefficient $e$, at FE level $L$,

$$(\dagger) \quad \mathcal{A}(\bar{D}^j, H^{-1}(D)) \xrightarrow{\text{Id}_L \otimes (\Delta_{e,L}^{-1})} \mathcal{A}(\bar{D}^j, V_L) \rightarrow \mathcal{A}(\bar{D}^{j-1}, V_L) \rightarrow \cdots$$

$$\mathcal{M}^{(j,1)}(r_M, f) \xrightarrow{\text{Id}_L \otimes (\Delta_{e,L}^{-1})} \mathcal{M}^{(j,1)}(r_M, v_{0,L}) \rightarrow \mathcal{M}^{(j-1,1)}(r_M, v_{1,L}) \rightarrow \cdots$$

$$\cdots \rightarrow V_L$$

$$\cdots \rightarrow \mathcal{M}^1(v_{j,L})$$

The alternative representation of the $(j, 1)$ mixed moment of $(r_M, f)$,

$$\mathcal{M}^{(j,1)}(r_M, f) = \sum_{1 \leq m_1 \leq M \leq M} \left( \int_D X_{m_1}(\omega) \cdots X_{m_j}(\omega) f(x, x) \, dP(\omega) \right) \phi_{m_1} \otimes \cdots \otimes \phi_{m_j} \in L^\infty(\bar{D}^j)$$

is processed by the diagram $(\dagger)$ and transformed into $\mathcal{M}^1(v_{j,L})$. The mean field computation algorithm consists therefore in running $(\dagger)$ for all $0 \leq j \leq n$ and with (1.6) as input. The total cost is shown to be essentially as high as the computational effort needed to run the first step in $(\dagger)$ for $j = n$ (definition of the following $j$ operators in $(\dagger)$ thus omitted). Taking into account the mixed moment symmetry w.r.t. permutations of the index set $(m_1, m_2, \ldots, m_j)$, we deduce that the algorithm requires solving as many deterministic diffusion problems in $D$ at FE discretization level $L$ as there are terms in the alternative representation (1.6) of the mixed moment of $(r_M, f)$, that is at least

$$M^n / n! \sim \varepsilon^{-O(1)} |\log \varepsilon| O((d-1)(|\log \varepsilon|) \text{ deterministic problems in } D \text{ with accuracy } \varepsilon.$$
as input in (\ref{eq:input}), the algorithm complexity is shown to become nearly optimal, consisting in solving
\[ \sim |\log \varepsilon| O(\log \varepsilon^{d/(d+1)}) < \varepsilon^{-\alpha(1)} \] deterministic problems in \( D \) with accuracy \( \varepsilon \). (1.7)

We conclude the presentation of the mean field computation algorithm by a comparison with the Monte Carlo method, in terms of overall complexity. Note first that sampling \( a, f \) over a finite sample set \( \Omega_0 \) requires knowledge of the data distribution everywhere in \( D \). Averaging the discrete (FE level \( L \)) solutions \( (u_L(\omega))_{\omega \in \Omega_0} \) of the sampled diffusion problems we obtain the accuracy estimate
\[ \| \mathcal{M}^1(u) - \frac{1}{|\Omega_0|} \sum_{\omega \in \Omega_0} u_L(\omega) \|_{H^1(D)} \sim \Phi(N_L) + \frac{1}{\sqrt{|\Omega_0|}}. \]

It follows that the sampling can not distinguish between a rough and a smooth (in the physical variable) fluctuation, and that the Monte Carlo method amounts in this case to solving
\[ \sim \varepsilon^{-2} \] deterministic problems in \( D \) with accuracy \( \varepsilon \), (1.8)

which is the desired quadratic complexity estimate.

Summarizing, in the first part of chapter 3 we propose a perturbation mean field computation algorithm and prove that additional information on the data (smoothness in the physical variable) allows a nearly optimal complexity estimate, asymptotically as \( \varepsilon \to 0 \). The algorithm is thus shown to be less expensive than the Monte Carlo simulation ((1.7) vs. (1.8)) and all other known algorithms (SG, W/gPC, FOSM etc.).

In the second part of chapter 3 we combine the perturbation algorithm presented above for the mean field computation with the sparse grid solution algorithm for the higher order moment problem discussed in chapter 2. The result is a new algorithm also of nearly optimal complexity for the computation of \( \mathcal{M}^k(u) \) for \( k \geq 2 \).

1.3 Notations and Definitions

We conclude this introductory part by setting up basic notations and definitions that will be constantly used throughout this work. We consider a bounded open set \( D \subset \mathbb{R}^d \) (the physical domain) with Lipschitz boundary \( \Gamma = \partial D \), we denote by \( \Lambda \) the Lebesgue completion of the Borelian \( \sigma \)-algebra on \( D \), and use the standard notation \( \lambda \) for the Lebesgue measure on \( D \). We model the uncertainty in the data through a probability space which we denote by \( (\Omega, \Sigma, P) \). Equipped with the product \( \sigma \)-algebra \( \Lambda \times \Sigma \), the space \( D \times \Omega \) becomes a \( \sigma \)-algebra.

Definition 1.3. If \( n \in \mathbb{N}_+ \), a measurable function \( g : D \times \Omega \to \mathbb{R}^n \) is called a random vector field on \( D \). If \( n = 1 \), \( g \) is simply a random field on \( D \).
1.3. NOTATIONS AND DEFINITIONS

In the following, mathematical objects (sets, functions, spaces, etc.) defined in terms of the physical domain $D$ will be referred to as deterministic, whereas those depending on $\Omega$ will be termed stochastic.

To formulate a minimal regularity requirement on the random fields we define the Bochner spaces of vector-valued measurable functions.

**Definition 1.4.** For an arbitrary Banach space $(B, \| \cdot \|_B)$, we define the space $L^0_{\text{step}}(\Omega, B)$ of all $B$-valued $\Sigma$-measurable step-functions on $\Omega$ by

$$L^0_{\text{step}}(\Omega, B) := \{ g : \Omega \to B \mid \exists (\Omega_i)_{i \in \mathbb{N}^+} \subset \Sigma \text{ partition of } \Omega, (b_i)_{i \in \mathbb{N}^+} \subset B \text{ s.t. } g = \sum_{i \in \mathbb{N}^+} \chi_{\Omega_i} b_i \},$$

where $\chi_{\Omega_i}$ denotes for $i \in \mathbb{N}^+$ the indicator function of the set $\Omega_i \subset \Omega$.

The Bochner space of $B$-valued $\Sigma$-measurable functions on $\Omega$ is then given by

$$L^0(\Omega, B) := \{ g : \Omega \to B \mid \exists (g_n)_{n \in \mathbb{N}^+} \subset L^0_{\text{step}}(\Omega, B) \text{ s.t. } g_n \xrightarrow{n \to \infty} g \text{ P-a.e. in } \Omega \}.$$

Clearly then, if $(B, \| \cdot \|)$ is a space of measurable (vector valued) functions on $D$, the corresponding Bochner space consists of random (vector) fields.

The notion of positivity for a stochastic diffusion coefficient reads

**Definition 1.5.** A random field $a \in L^0(\Omega, L^\infty(D))$ is called strictly positive if $a(\cdot)$ is $P$-a.e. strictly positive, i.e. there exists $N \in \Sigma$ such that $P(N) = 0$ and

$$\forall \omega \in \Omega \setminus N \exists a_{-,\omega}, a_{+,\omega} \in \mathbb{R}_+ \text{ s.t. } a_{-,\omega} \leq a(\omega) \leq a_{+,\omega} \text{ } \lambda \text{-a.e. in } D. \quad (1.9)$$

The existence of a unique stochastic solution to an elliptic problem with stochastic data follows then from the well-posedness of the corresponding deterministic problem (see also [Bab61]). In other words, the well-posedness is preserved by tensoring with measurable functions.

**Proposition 1.6.** If $a \in L^0(\Omega, L^\infty(D))$ is strictly positive and $f \in L^0(\Omega, H^{-1}(D))$, then there exists a unique $u \in L^0(\Omega, H^1_0(D))$ such that

$$-\text{div}(a(\omega) \nabla u(\omega)) = f(\omega) \text{ in } H^{-1}(D) \quad P\text{-a.e. in } \Omega. \quad (1.10)$$

**Proof.** Since the uniqueness part follows immediately from the well-posedness of the deterministic diffusion problem ((1.10) pointwise in $\Omega$), we only check here the existence of a (measurable) solution to (1.10). To this end, we consider two sequences $(a_n)_{n \geq 1}$, $(f_n)_{n \geq 1}$ as in Definition 1.4, as well as the corresponding representations $a_n = \sum_{i \in \mathbb{N}^+} \chi_{\Omega_i} a_{n,i}$ and $f_n = \sum_{i \in \mathbb{N}^+} \chi_{\Omega_i} f_{n,i}$ respectively (w.l.o.g. we assume them subordinate to the same partition of $\Omega$). With the notations in (1.9) we set for any $n \in \mathbb{N}^+$

$$I_n := \{ i \in \mathbb{N}^+ \mid a_{-,\omega}/2 \leq a_{n,i} \leq 2a_{+,\omega} \text{ } \lambda\text{-a.e. in } D \text{ for some } \omega \in \Omega_{n,i} \}. \quad (1.11)$$
From (1.9) and the $P$-a.e. convergence of the sequence $(a_n)_{n \in \mathbb{N}^+}$ to a we deduce that $\mathcal{I}_n \neq \emptyset$ for $n$ large enough. Clearly, for any $i \in \mathcal{I}_n$ the diffusion problem in $D$ with data $(a_{n,i}, f_{n,i})$ has a unique solution $u_{n,i} \in H^1_0(D)$. We then set
\[
 u_n := \sum_{i \in \mathcal{I}_n} \chi_{\Omega_{n,i}} \cdot u_{n,i} + \chi_{\Omega \setminus \bigcup_{i \in \mathcal{I}_n} \Omega_{n,i}} \cdot 0 \in L^0_{\text{step}}(\Omega, H^1_0(D)) \tag{1.12}
\]
and prove next that $(u_n)_{n \in \mathbb{N}^+}$ converges $P$-a.e. in $H^1_0(D)$. The limit, denoted by $u$, is in view of Definition 1.4 an element of $L^0(\Omega, H^1_0(D))$ and will be seen to satisfy (1.10).

Let $N \in \Sigma$ be a null set such that the pointwise convergence of the sequences $(a_n)_{n \geq 1}$, $(f_n)_{n \geq 1}$ as well as (1.9) hold on $\Omega \setminus N$. Fixing an arbitrary $\omega \in \Omega \setminus N$, it follows that there exists $n_\omega \geq 1$ such that for any $n \geq n_\omega$ there exists $i_\omega \in \mathcal{I}_n$ with the property $\omega \in \Omega_{n,i_\omega}$. This implies for any $n \geq n_\omega$ that $u_n(\omega) = u_{n,i_\omega}$ solves the deterministic diffusion problem in $D$ with data $(a_{n,i_\omega}, f_{n,i_\omega})$. The ellipticity condition (1.11), uniform in $n \geq n_\omega$, of all these problems ensures via a standard Strang estimate the $H^1_0(D)$ convergence of $u_n(\omega)$ to $u(\omega)$, which concludes the proof.

**Remark 1.7.** Under the assumptions in Proposition 1.6, the Doob-Dynkin Lemma (see e.g. [Fed69]) ensures the existence of a Borel measurable mapping
\[
h : \text{Ran} a^{L^\infty(D)} \times H^{-1}(D) \to H^1_0(D) \tag{1.13}
\]
such that
\[
u(\omega) = h(a(\omega), f(\omega)) \quad P\text{-a.e. in } \Omega. \tag{1.14}
\]
Throughout this work we assume that problem (1.10) satisfies the ellipticity condition (1.9) uniformly in $\Omega$, that is it holds

**Assumption 1.8.** There exist $a_-, a_+ \in \mathbb{R}_+$ such that
\[
0 < a_- \leq a(\omega) \leq a_+ \quad \lambda\text{-a.e. in } D, P\text{-a.e. in } \Omega. \tag{1.15}
\]
Note that for numerous engineering applications assumption (1.15) is satisfied (see e.g. [GS91]). However, (1.15) is clearly less general than (1.9) and does not cover e.g. the case of a log-normal diffusion coefficient, $a = \exp(Y)$ with $Y$ a gaussian random field (for details and applications of log-normally distributed random fields in groundwater and pollution monitoring problems see [Gil87]).
Chapter 2

Stochastic Source Term

Throughout this section the diffusion coefficient is assumed to be deterministic, that is, $a$ depends only on the physical variable $x \in D$. (1.15) becomes in this case

$$0 < a_- \leq a(x) \leq a_+ < \infty \quad \lambda\text{-a.e. in } D. \quad (2.1)$$

The source term $f$ of (1.10) is a random field with finite first (or higher) order moments

$$f \in L^k(\Omega, H^{-1}(D)) \quad k \in \mathbb{N}_+. \quad (2.2)$$

Under these assumptions we study the existence and computability of second and higher order moments of $u$ solution to (1.10), if the corresponding moments of the data $f$ are known.

2.1 Generalities

We begin with an introduction of the appropriate functional spaces which will help us define the moments associated to a random field. The standard reference for the material we briefly review in the following is [Yos95].

Definition 2.1. If $k \geq 1$ is an integer and $H$ is an arbitrary separable Hilbert space, we define the space of $L^k$, $H$-valued Bochner integrable functions (see e.g. [Yos95]) on $\Omega$ by

$$L^k(\Omega; H) := \left\{ f : \Omega \to H \mid f \text{ measurable}, \int_{\Omega} \|f(\omega)\|_H^k \, dP(\omega) < \infty \right\} / \sim$$

$$\|f\|_{L^k(\Omega; H)}^k := \int_{\Omega} \|f(\omega)\|_H^k \, dP(\omega),$$

where we use the same notation for a $P$-a.e. equivalence (denoted by $\sim$) class and one of its members.

A useful characterisation of these spaces follows from
Theorem 2.2 (Bochner). \( f \in L^k(\Omega; H) \) if and only if there exists a sequence of \( H \)-valued step functions \((f_j)_{j \in \mathbb{N}}\) such that

\[
f_j \to f \quad \text{P-a.e. on } \Omega \quad \text{and} \quad \int_{\Omega} \|f_j - f\|_H^k \to 0, \quad \text{as } j \to \infty. \tag{2.3}
\]

The case \( k = 1 \) is particularly important, since the elements of \( L^1(\Omega; H) \) can be integrated.

Theorem 2.3. If \( f \in L^1(\Omega; H) \), then the vector-valued integral

\[
\int_{\Omega} f(\omega) \, dP(\omega) \in H
\]

is well-defined by means of a sequence of \( H \)-valued step functions \((f_j)_{j \in \mathbb{N}}\) satisfying \((2.3)\) for \( k = 1 \), by

\[
\int_{\Omega} f(\omega) \, dP(\omega) := \lim_{j \to \infty} \int_{\Omega} f_j(\omega) \, dP(\omega) \in H. \tag{2.5}
\]

The role of the abstract Hilbert space \( H \) will be played in the context of the moments associated to random fields by the anisotropic Sobolev spaces.

Definition 2.4. If \( k \in \mathbb{N}^* \) and \( \mathbf{v} = (v_1, v_2, \ldots, v_k) \in \mathbb{R}^k \), we denote by

\[
H^\mathbf{v}(D^k) := \bigotimes_{j=1}^k H^{v_j}(D)
\]

the anisotropic Sobolev space of order \( \mathbf{v} \) on \( D^k \). Similarly we define the homogeneous anisotropic Sobolev space \( H^\mathbf{0}_0(D^k) \) (Convention: for any \( s \in \mathbb{R} \), \( s := (s, s, \ldots, s) \in \mathbb{R}^k \)).

Definition 2.5. If \( k \geq 1 \) is an integer, \( s \in \mathbb{R} \) and the random field \( f \) on \( D \) satisfies

\[
f \in L^k(\Omega; H^s(D)),
\]

so that the \( k \)-th order moment \( \mathcal{M}^k(f) \) of the random field \( f \) (sometimes called \( k \)-point correlation) is defined via Theorem 2.3 by

\[
\mathcal{M}^k(f) := \int_{\Omega} f(\omega) \otimes f(\omega) \otimes \cdots \otimes f(\omega) \, dP(\omega) \in H^s(D^k). \tag{2.8}
\]

Remark 2.6. Under the assumptions in Definition 2.5 it is easy to see that if \( s \geq 0 \), then the \( k \)-th order moment of \( f \) given by \((2.8)\) satisfies

\[
\mathcal{M}^k(f)(x_1, \ldots, x_k) = \int_{\Omega} f(x_1, \omega) \cdot f(x_2, \omega) \cdots f(x_k, \omega) \, dP(\omega), \tag{2.9}
\]

for any \((x_1, x_2, \ldots, x_k) \in D^k\).
2.2. THE MOMENT EQUATION

2.2 The Moment Equation

Here we derive a deterministic PDE for the $k$-th order moment $M^k(u)$, establish its well-posedness in anisotropic Sobolev scale in $D^k$, and study the solution regularity.

**Theorem 2.7.** If $f \in L^k(\Omega, H^{-1}(D))$, then $u$ solution to (1.10) satisfies

$$u \in L^k(\Omega, H^1_0(D)),$$

so that $M^k(u) \in H^1_0(D^k)$.

**Proof.** The conclusion follows at once from Definition 2.5 and the well-posedness of the deterministic diffusion problem in $D$,

$$||u(\omega)||_{H^1(D)} \leq c_a ||f(\omega)||_{H^{-1}(D)} \quad P\text{-a.e. } \omega \in \Omega,$$  

(2.11)

where the constant $c_a > 0$ depends only on the diffusion coefficient $a$.

We prove next that $M^k(u)$ satisfies the deterministic equation in $D^k$ obtained by tensorising (1.10) with itself $k$ times. To this end we introduce the following notations,

$$a^\otimes,k := \bigotimes_{j=1}^k a \in L^\infty(D^k)$$

$$\nabla^\otimes,k := \bigotimes_{j=1}^k \nabla \in B(H^1(D^k), \otimes_{j=1}^k L^2(D)^d)$$

where we denote by $B(X,Y)$ the space of bounded linear operators between the Hilbert spaces $X$ and $Y$, with $B(X) := B(X,X)$.

**Theorem 2.8.** $M^k(u)$ is the unique solution in $H^1_0(D^k)$ of the variational problem

$$\langle a^\otimes,k \nabla^\otimes,k M^k(u), \nabla^\otimes,k M \rangle_{L^2(D^k)} = \langle M^k(f), M \rangle_{H^{-1}(D^k),H^1_0(D^k)} \quad \forall M \in H^1_0(D^k).$$  

(2.12)

**Proof.** The existence and uniqueness of a solution to (2.12) are easily proved using the Lax-Milgram Lemma in appropriate anisotropic Sobolev spaces, noting that tensor products of bounded positive homeomorphisms between Hilbert spaces induce homeomorphisms between corresponding tensor products spaces.

Now, since $f \in L^k(\Omega, H^{-1}(D))$, there exists a sequence $(f_n)_{n \in \mathbb{N}}, (h_n)_{n \in \mathbb{N}}$ of $H$-valued step functions on $\Omega$ satisfying (2.3) with $H := H^{-1}(D)$. By dominated convergence and (2.5),

$$\lim_{n \to \infty} M^{\alpha}(f_n) = M^{\alpha}(f) \quad \text{in } H^{-1}(D^k).$$  

(2.13)

We write $f_n = \sum_{q \in J_n} f^q_n 1_{\Omega_{q,n}}$ where $1_{\Omega_{q,n}}$ denotes the indicator function of $\Omega_{q,n} \in \Sigma$ and, $f^q_n \in H^{-1}(D)$ for any $q,n$ (for each $n$ the family $(\Omega_{q,n})_{q \in J_n}$ is a partition of $\Omega$). To the
deterministic data $f^q_n$, we associate the solution $u^q_n \in H^1_0(D)$ of the corresponding diffusion problem,
\[
\langle a \nabla u^q_n, \nabla v \rangle_{L^2(D)^d} = \langle f^q_n, v \rangle_{H^{-1}(D), H^1_0(D)} \quad \forall v \in H^1_0(D),
\]
and set $u_n := \sum_{q \in I_n} u^q_n \mathbf{1}_{\Omega_{q,n}}$. The continuous dependence (2.11) on the data of the solution of the deterministic problem and (2.3) for $f$ imply
\[
\lim_{n \to \infty} u_n \to u \quad P\text{-a.e. on } \Omega, \quad \lim_{n \to \infty} \int_{\Omega} \|u(\omega) - u_n(\omega)\|_{H^1_0(D)}^2 \, dP(\omega) = 0.
\]
From Definition 2.5, (2.15) and (2.5) we deduce that
\[
\lim_{n \to \infty} \mathcal{M}^k(u_n) = \mathcal{M}^k(u) \quad \text{in } H^1(D).
\]
Choosing in (2.14) $k$ different deterministic test functions $v_1, v_2, \ldots, v_k$, taking the product of the resulting $k$ equations and summing over $q$ with weights $P(\Omega_{q,n})$, we obtain that
\[
\mathcal{M}^k(u_n) \text{ solves the deterministic problem}
\]
\[
\langle a^{\otimes k} \nabla^{\otimes k} \mathcal{M}^k(u_n), \nabla^{\otimes k} \mathcal{M} \rangle_{L^2(D)^{dk}} = \langle \mathcal{M}^k(f_n), \mathcal{M} \rangle_{H^{-1}(D^k), H^1_0(D^k)} \quad \forall \mathcal{M} \in H^1_0(D^k)
\]
(“use here that tensor products of total sets in Hilbert spaces are total in product spaces). The desired equation for $\mathcal{M}^k(u)$ follows then from (2.13) and (2.16) by letting $n \to \infty$ in (2.17).

The regularity of $\mathcal{M}^k(u)$ follows naturally from that of the data $\mathcal{M}^\alpha(f)$, and the result, as well as its proof, is analogous to the one presented in [ST03a] for $k = 2$ (bounded invertibility of operators acting in Hilbert spaces is preserved by tensorization). We only state it, as follows. Recall that a deterministic diffusion problem in $D$ is said to satisfy the shift theorem at order $s > 0$ if $H^{-1+s}$ regularity of the data implies $H^{1+s}$ regularity of the solution.

**Theorem 2.9.** If the deterministic diffusion problem in $D$ with coefficient $a$ satisfies the shift theorem at order $s$, then (2.12) satisfies a shift theorem at order $s$ too, in the sense that $\mathcal{M}^k(f) \in H^{-1+s}(D^k)$ implies $\mathcal{M}^k(u) \in H^{1+s}(D^k)$.

**Remark 2.10.** In the case of a polygon or polyhedron $D$, a shift theorem at order $s \geq 0$ holds in weighted spaces $H^{1+s,2}_0(D)$ (see [BG88]). The proof of Theorem 2.9 can be correspondingly adapted to deduce then a shift theorem for the moment equation (2.12) in an anisotropic weighted Sobolev scale.

**2.3 Discretization with Sparse FEM**

We investigate now the numerical approximation of $\mathcal{M}^k(u)$, using the Finite Element Method for the deterministic elliptic problem (2.12). We start by defining general FE spaces, so let
\[
V_0 \subset V_1 \subset \ldots \subset V_L \subset \ldots \subset H^1_0(D)
\]
be a dense hierarchical sequence of finite dimensional subspaces of $H_0^1(D)$, with $N_L := \dim(V_L) < \infty$ for all $L$.

**Assumption 2.11.** The following approximation property holds,

$$\min_{v \in V_L} ||u - v||_{H_0^1(D)} \leq \Phi(N_L, s)||u||_{H^{s+1}(D)}, \quad \forall u \in H^{s+1}(D) \cap H_0^1(D), \quad (2.19)$$

where $\Phi(N, s) \to 0$ for $s > 0$ as $N \to \infty$ denotes the convergence rate of the given FE space sequence.

For regular solutions the usual FE spaces based on quasiuniform, shape regular meshes are suitable.

**Example 2.12.** If $T = \{T_L\}_{L \in \mathbb{N}}$ is a nested sequence of regular triangulations of $D$ of meshwidth $h_L = h_{L-1}/2$, we choose $V_L$ to be the space of all continuous piecewise polynomials of degree $p$ on $T_L$ vanishing on $\partial D$. Then Assumption 2.11 holds with

$$N_L \sim c_L 2^{dL}, \quad \Phi(N, s) = c_L N^{-\delta}, \quad \delta := \min\{p, s\}/d. \quad (2.20)$$

Since the $k$-th order moment $\mathcal{M}^k(u)$ of $u$ solves the elliptic problem (2.12) on $D^k$, we construct FE spaces in $D^k$ starting from the FE spaces $\{V_L\}_{L \geq 0}$ in $D$. Full tensor product spaces,

$$V_L := \bigotimes_{j=1}^k V_L$$

are natural candidates, but, due to efficiency reasons, we use here the sparse tensor product spaces that are defined by (see [Zen91], [BG99])

$$\hat{V}_L := \text{Span} \left\{ \bigotimes_{j=1}^k V_{i_j} \mid 0 \leq i_1 + i_2 + \ldots + i_k \leq L \right\}. \quad (2.21)$$

Another description of the sparse tensor space (2.21) can be given in terms of the hierarchic excess $W_L$ at level $L \geq 0$ of the scale $\{V_L\}_{L \geq 0}$, which is chosen to be an arbitrary algebraic complement of $V_{L-1}$ in $V_L$ (set $V_{-1} := \{0\}$). Since then

$$V_L = \bigoplus_{0 \leq i \leq L} W_i, \quad (2.22)$$

it is easily seen that $\hat{V}_L$ admits the direct algebraic (not necessarily $L^2$ or $H^1$ orthogonal!) decomposition

$$\hat{V}_L := \bigoplus_{0 \leq i_1 + i_2 + \ldots + i_k \leq L} \bigotimes_{j=1}^k W_{i_j} \subset \bigoplus_{0 \leq i_1, i_2, \ldots, i_k \leq L} \bigotimes_{j=1}^k W_{i_j} = \bigotimes_{j=1}^k V_L. \quad (2.23)$$

The discretized version of (2.12) using the sparse FE space $\hat{V}_L$ reads

$$\langle a^{\otimes k} \nabla^{\otimes k} \mathcal{M}^k_L(u), \nabla^{\otimes k} \mathcal{M}^k_L \rangle_{L^2(D)} = \langle \mathcal{M}^k(f), \mathcal{M}^k_L \rangle_{H^{-1}(D^k), H_0^1(D^k)} \quad \forall \mathcal{M}_L \in \hat{V}_L, \quad (2.24)$$
where we denote by $M^k_L(u) \in \tilde{V}_L$ the discrete solution of (2.12). The approximation property (2.19) allows an estimate of the discretization error in terms of the functional $\Phi$, as follows.

**Proposition 2.13.** If $M^k_L(u)$ is the solution to (2.24), $L \geq k - 1$ and the approximation property (2.19) holds, then

$$
\|M^k(u) - M^k_L(u)\|_{H^1(D^k)}^2 \leq c_{a,k} \sum_{J \subseteq \{1, \ldots, k\}} \sum_{\text{Card}(J) = j} \|M^k(u)\|_{H^{2-k+j+1}(D^k)}^2 
$$

(2.25)

where $e_j \in \{0, 1\}^k$, $e_j(j) = 1$ iff $j \in J$ and

$$
c(j, \Phi) = \sum_{m=1}^{j-1} \sum_{t=k}^{L-m+1} (\Phi(N_{l_1}, s) \cdot \Phi(N_{l_2}, s) \cdots \Phi(N_{l_m}, s) \cdot \Phi(N_0, s))^2 + \sum_{l_1+\ldots+l_j=L-j+1} \Phi(N_{l_1}, s)^2 \cdot \Phi(N_{l_2}, s)^2 \cdots \Phi(N_{l_j}, s)^2 
$$

(2.26)

Besides, the constant $c_{a,k}$ depends exponentially on $k$, that is, $c_{a,k} = c^k_a$.

**Proof.** The estimate (2.25) follows using the quasi-optimality of the FE solution to (2.24), the approximation property (2.19) and the description (2.23) of the sparse tensor space with $W_L$ defined as the orthogonal complement of $V_{L-1}$ in $V_L$ w.r.t the standard scalar product $\langle \cdot, \cdot \rangle$ in $H^1_0(D)$. More precisely, we consider the following orthogonal decomposition in $H^1_0(D^k)$ equipped with the tensor product Hilbert structure induced by $\otimes_{i=1}^k \langle \cdot, \cdot \rangle$ (here and in the following orthogonality in $H^1_0(D^k)$ is to be understood w.r.t. this Hilbert structure).

$$
M^k(u) - P^k_{S_L}(M^k(u)) = \sum_{\alpha_1 + \alpha_2 + \ldots + \alpha_k \geq L+1} \otimes_{i=1}^k P^l_{\alpha_i} M^k(u) 
$$

(2.27)

where $P^l_{\alpha}$ denotes the orthogonal projection on $W_\alpha$ w.r.t. $\langle \cdot, \cdot \rangle$, acting in the $i$-th dimension of $D^k$. As the notation suggests, $P^k_{\tilde{V}_L}$ denotes the $H^1_0(D^k)$-orthogonal projection on $\tilde{V}_L$. Further we consider $Q^l_{\alpha}$ to be the projection on $V_\alpha$ acting in the $i$-th direction of $D^k$ and note that the sum in the decomposition (2.27) is $H^1_0(D^k)$-orthogonal, since the excesses $(W_L)_{\alpha \in \mathbb{N}}$ are pairwise $H^1_0(D)$-orthogonal. We rewrite the r.h.s. of (2.27), pointing out those dimensions $1 \leq j \leq k$ for which $\alpha_j = 0$ (coarsest approximation),

$$
\sum_{\alpha_1 + \alpha_2 + \ldots + \alpha_k \geq L+1} \left( \otimes_{i=1}^k P^l_{\alpha_i} \right) M^k(u) = \sum_{p=1}^k \sum_{J \subseteq \{1, \ldots, k\}} \sum_{\alpha_j \geq 1} \left( \otimes_{j \in J} P^j_{\alpha_j} \otimes_{j \notin J} P^j_0 \right) M^k(u),
$$
and we cast the first inner sum of projections for $J = \{j_1, j_2, \ldots, j_p\}$ in the form

$$
\sum \sum_{\sum_{n=1}^{p} \alpha_n \geq L+1} P_{\alpha_n}^{j_n} \otimes P_0^j = (\text{Id} - Q_L^j) \otimes (\text{Id} - P_0^{j_n}) \otimes P_0^j
$$

$$
+ \sum_{\sum_{n=1}^{p} \alpha_n \geq L+1} P_{\alpha_n}^{j_n} \otimes (\text{Id} - Q_L^{j_n}) \otimes (\text{Id} - P_0^{j_n}) \otimes P_0^j
$$

$$
+ \sum_{\sum_{n=1}^{p} \alpha_n \geq L+1} P_{\alpha_n}^{j_n} \otimes (\text{Id} - Q_L^{j_n}) \otimes (\text{Id} - P_0^{j_n}) \otimes P_0^j
$$

$$
+ \ldots + \sum_{\sum_{n=1}^{p} \alpha_n \geq L+1} P_{\alpha_n}^{j_n} \otimes (\text{Id} - Q_L^{j_n}) \otimes (\text{Id} - P_0^{j_n}) \otimes P_0^j.
$$

Here the $l$-th sum in the decomposition (2.28) consists of those terms in the l.h.s. corresponding to indices $\alpha_1, \alpha_2, \ldots, \alpha_p \geq 1$ with $\sum_{n=1}^{p} \alpha_n \geq L+1$ and for which $l \leq p$ is the smallest integer with the property $\sum_{n=1}^{l} \alpha_n \geq L+1$.

The conclusion follows from (2.28) using the trivial estimate $\|P_\alpha\| \leq \|\text{Id} - Q_{\alpha-1}\|$ (operator norm in $H^0(D^k)$) and the approximation property (2.19).

Specializing Proposition 2.13 by choosing the FE spaces from Example 2.12 we obtain

**Corollary 2.14.** For the sparse tensor product FE space based on the construction given in Example 2.12 and for any $L \geq k-1$ it holds

$$
\|M^k(u) - M_L^k(u)\|_{H^1(D^k)} \leq c_{a,k,T} (\log N_L)^{(k-1)/2} N_L^{-\delta} \cdot \|M^k(u)\|_{H^{s+1}(D^k)}
$$

$$
= c_{a,k,T,u} (\log N_L)^{(k-1)/2} N_L^{-\delta}
$$

(2.29)

and

$$
\dim \hat{V}_L \leq c_{k,T} (\log N_L)^{k-1} N_L,
$$

(2.30)

where $s = (s, s, \ldots, s)$, $\delta = \min\{p, s\}/d$, and all the constants depend exponentially on $k$.

Note that the full tensor product space would require $O(N_L^k)$ degrees of freedom for a relative tolerance $O(N_L^{-\delta})$.

**2.4 Iterative Solution and Complexity**

Sparse FE spaces discussed in section 2.3 allow an important reduction of the number of degrees of freedom needed to compute a discrete solution to the moment problem (2.12) up to a prescribed accuracy. After choosing a basis in the sparse FE space $\hat{V}_L$ and computing the corresponding stiffness matrix $\hat{S}_L^k$, finding the discrete solution of (2.24) amounts to solving the linear system

$$
\hat{S}_L^k \hat{M}^k(u) = \hat{M}^k(f).
$$

(2.31)
In order to be able to solve (2.31) efficiently using an iterative method (e.g. the conjugate gradient (CG) method), we must ensure well-conditioning and sparsity of the involved matrix $\hat{S}^L$. While the first property can be ensured by a wavelet preconditioning procedure, the latter (which does not hold, in fact) will be replaced by a proper use of the anistropic structure of the problem, which will show that computation and storage of $\hat{S}^L$ is not necessary.

Concerning notations, $\mathcal{F}$ will be in the following a family of double indices running in $\mathbb{N}^d \times \mathbb{N}^d$.

**Assumption 2.15.** There exist a family $\Psi = (\psi_{j,i})_{(j,i) \in \mathcal{F}} \subset H_0^1(D)$ and constants $c_1, c_2 > 0$ such that any $u \in H_0^1(D)$ can be expanded as a convergent series,

$$u = \sum_{(j,i) \in \mathcal{F}} c_{j,i} \psi_{j,i} \in H_0^1(D)$$

and the following stability condition is fulfilled

$$c_{1,\psi} \sum_{(j,i) \in \mathcal{F}} |c_{j,i}|^2 \leq \| \sum_{(j,i) \in \mathcal{F}} c_{j,i} \psi_{j,i} \|_{H_0^1(D)}^2 \leq c_{2,\psi} \sum_{(j,i) \in \mathcal{F}} |c_{j,i}|^2.$$  \hspace{1cm} (2.33)

Constructions of families $\Psi$ satisfying Assumption 2.15 are known for intervals ($D = ]0,1[)$, hypercubes ($D = ]0,1[^d$) or polygonal domains (see e.g. [DS99]).

**Example 2.16 (hierarchical basis on interval).** For $D = ]0,1[$, let $\phi$ be the hat function on $\mathbb{R}$, piecewise linear, taking values 0, 1, 0 at 0, 1/2, 1 and vanishing outside $]0,1[$. Setting $\mathcal{F} := \{(j,i) \mid 0 \leq j, 1 \leq i \leq 2^j\}$ and

$$\psi_{j,i}(x) := 2^{-j/2} \phi(2^j x - i + 1) \ \forall x \in ]0,1[,$$

the family $\Psi = (\psi_{j,i})_{(j,i) \in \mathcal{F}}$ satisfies Assumption 2.15.

**Example 2.17 (prowavelets on interval).** With $D$, $\mathcal{F}$ and $\phi$ as above, we define on $\mathbb{R}$ the function $\psi$, piecewise linear, taking values (1, -6, 10, -6, 1) at (1/2, 1, 3/2, 2, 5/2) and vanishing outside $]0,3[$. Similarly, $\psi^l$ take (9, -6, 1) at (1/2, 1, 3/2) and $\psi^r$ assumes values (1, -6, 9) at (1/2, 1, 3/2). Further, we define the scaling function by $\psi_{0,1} := \phi$ and the boundary wavelets,

$$\psi_{j,1}(x) := 2^{-j/2} \psi^l(2^j x)$$

$$\psi_{j,2}(x) := 2^{-j/2} \psi^r(2^j x - 2^j + 1) \ \forall x \in ]0,1[ , \forall j \geq 1.$$  \hspace{1cm} (2.34)

The interior wavelets are given by

$$\psi_{j,i}(x) := 2^{-j/2} \psi(2^j x - i + 2) \ \forall x \in ]0,1[ , \forall 2 \leq i \leq 2^j - 1 , j \geq 2.$$  \hspace{1cm} (2.34)

The family $\Psi = (\psi_{j,i})_{(j,i) \in \mathcal{F}}$ satisfies then Assumption 2.15.
For further examples we refer the reader to [Coh03], [Dah97] and the references therein.

**Example 2.18 (wavelets on hypercube).** In $D = [0, 1]^d$, choose $\mathcal{F} := \{(j, i) \in \mathbb{N}^d \times \mathbb{N}^d \mid 0 \leq j, 1 \leq i \leq 2^j\}$ (inequalities involving multi-indices are understood componentwise). Tensorizing the family constructed in Example 2.17 by setting

$$\psi_{j,i}(x) = \prod_{q=1}^{d} \psi_{j_q,i_q} (x_q) \quad \forall x = (x_q)_{1 \leq q \leq d} \in D$$

we obtain (after rescaling in $H^1(D)$) a family $\Psi = (\psi_{j,i})_{(j,i) \in \mathcal{F}}$ which satisfies Assumption 2.15 (see e.g. [GO95]).

Formally, an increasing $\text{FE}$ space sequence in $D \subset \mathbb{R}^d$ can be defined in terms of the family $\Psi = (\psi_{j,i})_{(j,i) \in \mathcal{F}}$ by

$$V_L := \text{Span}\{\psi_{j,i} \mid 0 \leq |j|_\infty \leq L\}$$

($j$ is in general a multi-index, so that $|j|_\infty := \max_{1 \leq q \leq d} j_q$).

We consider the algebraic complement $W_L$ of $V_{L-1}$ in $V_L$ given by

$$W_L := \text{Span}\{\psi_{j,i} \mid |j|_\infty = L\},$$

and obtain, via (2.23), the following representation of the sparse tensor space $\hat{V}_L$,

$$\hat{V}_L = \text{Span}\{\psi_{j,i} := \bigotimes_{\nu=1}^{k} \psi_{j(\nu),i(\nu)} \mid \sum_{\nu=1}^{k} |j(\nu)|_\infty \leq L\},$$

where $j(\nu)$ denotes the $\nu$-th line of the $k \times d$ matrix $j$ and similarly for $i$.

The algebraic excess $\hat{W}_L$ of the sparse tensor scale $(\hat{V}_L)_{L \geq 0}$ is then given by

$$\hat{W}_L = \text{Span}\{\psi_{j,i} := \bigotimes_{\nu=1}^{k} \psi_{j(\nu),i(\nu)} \mid \sum_{\nu=1}^{k} |j(\nu)|_\infty = L\},$$

and can be further decomposed as

$$\hat{W}_L = \bigoplus_{\|l\|_\infty = L} W_l$$

with

$$W_l = \text{Span}\{\psi_{j,i} \mid |j(\nu)|_\infty = l_\nu\},$$

where

$$\|l\| := l_1 + l_2 + \ldots + l_k \quad \forall l \in \mathbb{N}^k.$$

For further use we collect, for $L \geq 0$, the basis functions in the definition (2.37) of $W_L$ in a vector denoted $\Psi_L$. Similarly, for $l \in \mathbb{N}^k$ we consider $\Psi_l$ to be the vector containing all basis functions of $W_l$ as defined in (2.40).

Concerning the properties of the stiffness matrix $\hat{S}^L$ that are of interest for solving (2.31) (well-conditioning and sparsity), it holds
Proposition 2.19. The matrices \((\tilde{S}^L)_L\\geq 0\) have uniformly bounded condition number
\[
\kappa(\tilde{S}^L) \leq c_{n,T,\pi,k} \quad \forall L \geq 0,
\]  
where the upper bound \(c_{n,T,\pi,k}\) depends exponentially on \(k\).

Proof. From (2.33) it follows that the basis \((\psi_{j,i})_{(j,i)\in F}\) gives a homeomorphism of Hilbert spaces between \(\ell^2\) and \(L^2_0(D)\), or that
\[
u = \sum_{(j,i)\in F} c_{j,i} \psi_{j,i} \quad \rightarrow \quad |u|^2_w := \sum_{(j,i)\in F} |c_{j,i}|^2
\]  
defines an equivalent norm on \(L^2_0(D)\). The same holds then for the basis \(\psi_{j,i}\) introduced in (2.38). It follows that for \(M := (M_{j,i,j',i'})_{j,i,j',i'} \in \mathbb{R}^{N_L^2}\) with \(\tilde{V}_L, M := \sum_{j,i} M_{j,i} \psi_{j,i}\) is an element of \(\tilde{V}_L\) and
\[
\langle \tilde{S}^L M, M \rangle_{\mathbb{R}^{N_L^2}} = \langle a^{\otimes,k} \nabla^{\otimes,k} M, \nabla^{\otimes,k} M \rangle_{L^2(D)^{4k}} \sim \|M\|^2_{L^2(D)^{4k}} \sim \sum_{j,i} |M_{j,i}|^2 = \|M\|^2_{\mathbb{R}^{N_L^2}}.
\]

Proposition 2.20. For Examples 2.16, 2.17, 2.18 above as well as for similar wavelet constructions the matrix \(\tilde{S}^L\) is not sparse, in the sense that (compare (2.30))
\[
nnz(\tilde{S}^L) \geq O(N_L^2) \quad \forall L \geq 0.
\]  
Proof. It is easily seen that the entries of \(\tilde{S}^L\) corresponding to the indices \(i,j,i',j'\) with \(j(1)=j'(2) = (L, L, \ldots, L)\) are in general nonzero, implying the desired lower bound.

The nonsparse makes the storage and use of \(\tilde{S}^L\) rather costly, and shows that ideas beyond multilevel preconditioning are needed for an efficient solution of (2.31). The key observation is the special structure of the discrete operator (or, equivalently, of \(\tilde{S}^L\)), which inherits (blockwise) the tensor product structure of the continuous operator (see (2.12)).

A matrix-vector multiplication algorithm of log-linear complexity will be derived in the following, exploiting the special tensor product structure of the problem. It will be shown that, in terms of memory, only (blockwise) storage of the matrix \(S^L\) corresponding to the case \(k = 1\) is necessary. Relating \(\tilde{S}^L\) to \(S^L\) will show that one step of CG algorithm can be performed by manipulating only sparse blocks of \(S^L\). Moreover, an algorithm of log-linear complexity in \(N_L\) for performing the matrix-vector multiplication required by the CG method will be derived.

Storage of the load vector is necessary too, but, due to (2.30), the corresponding memory requirement grows only log-linearly in \(N_L\), as \(L \to \infty\).

We start with the derivation of the relation between \(\tilde{S}^L\) and \(S^L\). To this end we denote by \(\langle \cdot, \cdot \rangle_w\) the scalar product associated with the norm (2.42). \(\langle \cdot, \cdot \rangle_w\) is obviously equivalent to
the standard scalar product in \( H^1_0(D) \) and \( (\psi_{j,i})_{(j,i) \in \mathcal{F}} \) is an orthonormal basis of \( H^1_0(D) \) equipped with \( \langle \cdot, \cdot \rangle_w \). Let \( P_L \) and \( Q_L \) be the orthogonal projections in \( H^1_0(D) \) w.r.t. \( \langle \cdot, \cdot \rangle_w \), on \( V_L \) and \( W_L \) given by (2.36) and (2.37) respectively. Clearly then,

\[
P_L = \sum_{l=0}^{L} Q_l.
\]

Correspondingly, we denote by \( \hat{P}_L \) and \( \hat{Q}_L \) the orthogonal projections on \( \hat{V}_L \) and \( \hat{W}_L \) (see (2.38), (2.39)) w.r.t. the scalar product on \( H^1_0(D^k) \) obtained by tensorization of \( \langle \cdot, \cdot \rangle_w \).

On account of (2.38), (2.39) we obtain the multilevel decomposition

\[
\hat{P}_L = \sum_{l=0}^{L} \hat{Q}_l,
\]

as well as the detail decomposition

\[
\hat{Q}_L = \sum_{|\ell| = L} Q_{\ell},
\]

where

\[
Q_{\ell} := \bigotimes_{\nu=1}^{k} Q_{\nu},
\]

is the projection on the space \( W_{\ell} \) introduced in (2.40).

We further denote by \( Q^k \) the bilinear form of the moment problem (2.12),

\[
Q^k := Q \otimes Q \otimes \ldots \otimes Q,
\]

where

\[
Q(u, v) := \langle A \nabla u, \nabla v \rangle_{L^2(D)} \quad \forall u, v \in H^1_0(D).
\]

The discrete problem in \( \hat{V}_L \) is then given by the bilinear form

\[
Q^k_L(u, v) := Q^k(\hat{P}_L u, \hat{P}_L v) \quad \forall u, v \in \hat{V}_L \subset H^1_0(D^k),
\]

or, inserting (2.44) and (2.45) in (2.49), by

\[
Q^k_L(u, v) = \sum_{l,l'=0}^{L} \sum_{|\ell| = |l|, |\ell'| = |l'|} Q^k(Q_{l} u, Q_{l'} v) \quad \forall u, v \in \hat{V}_L.
\]

Recalling that \( \Psi_{\ell} \) is the vector containing the basis functions of \( W_{\ell} \) given in (2.40), we write

\[
Q_{l} u = u_{\ell}^T \cdot \Psi_{\ell},
\]

with real vector coefficients \( u_{\ell} \) and similarly for \( v \).

Using (2.51) in (2.50), we obtain

\[
Q^k_L(u, v) = \sum_{l,l'=0}^{L} \sum_{|\ell| = |l|, |\ell'| = |l'|} u_{\ell}^T \cdot \hat{S}_{l,l'} \cdot v_{\ell'},
\]
where the matrix $\hat{S}_{LL}^{k}$ is given by evaluating the bilinear form on the basis functions,

$$\hat{S}_{LL}^{k} := Q^k(\Psi_l, \Psi_{l'})$$

But, in view of (2.47) and (2.39), we have

$$\hat{S}_{LL}^{k} = Q^k(\Psi_l, \Psi_{l'}) = \bigotimes_{\nu=1}^{k} Q(\Psi_{l_{\nu}}, \Psi_{l'_{\nu}}) = \bigotimes_{\nu=1}^{k} S_{l_{\nu}, l'_{\nu}}^{\nu}, \quad (2.53)$$

where

$$S_{l, l'}^{\nu} := Q(\Psi_l, \Psi_{l'}), \quad \forall 0 \leq l, l' \leq L \quad (2.54)$$

are the blocks of the stiffness matrix $S^{L}$ corresponding to the deterministic diffusion problem in $D$ with coefficient $a$ (or, equivalently, to the case $k = 1$).

The representation formulae (2.52) and (2.53) show that

$$Q_{L}^{k}(u, v) = \sum_{l, l'=0}^{L} \sum_{||=l, ||'=l'}^{\|\|=\|'} w_{l}^{T} \cdot \left( \bigotimes_{\nu=1}^{k} S_{l_{\nu}, l'_{\nu}}^{\nu} \right) \cdot w_{l'}, \quad (2.55)$$

that is the stiffness matrix $\hat{S}_{L}^{k}$ of the $k$-th moment problem computed w.r.t. the basis (2.38) of the FE space $\hat{V}_{L}$ has a block structure

$$\hat{S}_{L}^{k} = (\hat{S}_{LL}^{k})_{||=leN^k, ||'=leN^k}$$

and each block is a tensor product of certain blocks of the stiffness matrix $S^{L}$.

Moreover, $S^{L}$ is almost sparse, once for the basis $(\psi_{j,i})_{(i,j)\in\mathcal{I}}$ the following local support assumption holds true. We remark that Examples 2.16, 2.17, 2.18 as well as similar wavelet-type constructions are in this category.

**Assumption 2.21.** There exists $p \in \mathbb{N}_{+}$ such that for all $1 \leq i \leq 2^d \in \mathbb{N}^d$ and $j' \in \mathbb{N}^d$, the set $\text{supp}(\psi_{j,i}) \cap \text{supp}(\psi_{j',i'})$ has nonempty interior for at most $p^d \cdot \prod_{q=1}^{d} \max(1, 2^{d-j_{q}})$ values of $i'$.

**Remark 2.22.** From Assumption 2.21 it follows by a simple counting argument that

$$\text{nnz}(S_{l, l'}^{\nu}) \leq p^d \cdot (\min(l, l') + 1)^{d-1} \cdot 2^d \cdot \max(l, l') \quad \forall 0 \leq l, l' \leq L. \quad (2.56)$$

To formulate the matrix-vector multiplication algorithm we consider also, for each pair $l = (l_{\nu})_{\nu=1}^{k}, l' = (l'_{\nu})_{\nu=1}^{k}$, a reordering (permutation) $\sigma_{L}^{k}$ of $\{1, 2, \ldots, k\}$ such that

$$\sum_{\nu=1}^{q} l_{\sigma(\nu)} + \sum_{\nu=q+1}^{k} l'_{\sigma(\nu)} \leq \max \left\{ \sum_{\nu=1}^{k} l_{\nu}, \sum_{\nu=1}^{k} l'_{\nu} \right\} \quad \forall 1 \leq q \leq k. \quad (2.57)$$

The existence of such a permutation $\sigma$ is easy to prove, by choosing $x_{\nu} = l_{\nu}, y_{\nu} = l'_{\nu}$ for any $1 \leq \nu \leq k$ in the following result.
Lemma 2.23. If \((x_\nu)_{1\leq \nu \leq k}\) and \((y_\nu)_{1\leq \nu \leq k}\) are two families of positive real numbers, then there exists a permutation \(\sigma\) of the set \(\{1,2,\ldots,k\}\) such that

\[
\sum_{\nu=1}^{q} x_{\sigma^{(\nu)}} + \sum_{\nu=q+1}^{k} y_{\sigma^{(\nu)}} \leq \max \left\{ \sum_{\nu=1}^{k} x_\nu, \sum_{\nu=1}^{k} y_\nu \right\}, \quad \forall 1 \leq q \leq k.
\]

Proof. We use induction on \(k\). Since for \(k = 1\) the claim is trivial, we assume that it holds also for some \(k \geq 1\). Consider \((x_\nu)_{1\leq \nu \leq k+1}\) and \((y_\nu)_{1\leq \nu \leq k+1}\) two families of positive real numbers and define \(z_\nu := x_\nu\) for \(1 \leq \nu \leq k-1\) and \(z_k := x_k + x_{k+1}\), as well as \(t_\nu := y_\nu\) for \(1 \leq \nu \leq k-1\) and \(t_k := y_k + y_{k+1}\). The induction assumption ensures the existence of a permutation \(\tau\) of \(\{1,2,\ldots,k\}\) such that

\[
\sum_{\nu=1}^{q} x_{\tau^{(\nu)}} + \sum_{\nu=q+1}^{k} y_{\tau^{(\nu)}} \leq \max \left\{ \sum_{\nu=1}^{k} x_\nu, \sum_{\nu=1}^{k} y_\nu \right\}, \quad \forall 1 \leq q \leq k.
\]

We set then \(\sigma^{(\nu)} := \tau^{(\nu)}\) for all \(\nu < \tau^{-1}(k)\) and \(\sigma^{(\nu)} := \tau^{(\nu-1)}\) for all \(\nu > \tau^{-1}(k)+1\). Now, if \(y_k + x_{k+1} \leq x_k + y_{k+1}\) holds, we set \(\sigma(\tau^{-1}(k)) := k\), \(\sigma(\tau^{-1}(k)+1) := k+1\). Otherwise, that is if \(y_k + x_{k+1} > x_k + y_{k+1}\), we define \(\sigma(\tau^{-1}(k)) := k+1\) and \(\sigma(\tau^{-1}(k)+1) := k\). With this choice for \(\sigma\) one can easily check the inequalities (2.58). \(\square\)

In order to simplify the exposition of the algorithm we introduce, for any \(1 \leq q \leq k\), an arbitrary index pair \((l, l')\) and a permutation \(\sigma = \sigma_{l, l'}\) associated to it in the sense explained above, the following tensor product matrices

\[
T_{l, l', q}^L := \bigotimes_{\nu=1}^{k} U_\nu, \quad \text{where} \quad U_\nu := \begin{cases} \text{Id}_{l, l'}, & \nu \in \{\sigma(1), \sigma(2), \ldots, \sigma(q-1)\} \\ S_{l, l'}^{L_{\nu(q)}, \nu(q)}, & \nu = \sigma(q), \\ \text{Id}_{l, l'}, & \nu \in \{\sigma(q+1), \sigma(q+2), \ldots, \sigma(k)\} \end{cases}
\]

and \(\text{Id}_{l,l}\) denotes for \(l \geq 0\) the identity matrix of size \(\text{dim}W_l\). With these notations, each block in (2.55) can be expressed as a product of matrices of the type introduced in (2.60),

\[
\bigotimes_{\nu=1}^{k} S_{l, l'}^{L_{\nu}, L_{\nu+1}} = T_{l, l', q}^L \cdot T_{l, l', k-1}^L \cdots T_{l, l', 1}^L.
\]

For later use we remark that (2.56) implies the following sparsity estimate for \(T_{l, l', q}^L\).

Remark 2.24. It holds

\[
\text{nnz}(T_{l, l', q}^L) \leq c_{T,k,q} \prod_{\nu=1}^{q-1} (l_{\sigma^{(\nu)}} + 1)^d \cdot (\min(l_{\sigma^{(q)}}, l_{\sigma^{(q)}}) + 1)^{d-1} \cdot \prod_{\nu=q+1}^{k} (l_{\sigma^{(\nu)}} + 1)^d 
\cdot 2^d (\sum_{\nu=1}^{q-1} i_{\sigma^{(\nu)}} + \max(l_{\sigma^{(q)}}, l_{\sigma^{(q)})} + \sum_{\nu=q+1}^{k} i_{\sigma^{(\nu)}}). \quad (2.63)
\]
Proof. The estimate (2.63) follows from the obvious equality
\[
\text{nnz}(T_{L\ell,\mu}^L) = \prod_{q=1}^{\mu-1} \dim W_{i(q)} \cdot \text{nnz}(S_{i(k),\ell'}^L) \cdot \prod_{q=\mu+1}^{K} \dim W_{i(q)},
\]
the asymptotic estimate \(\dim W_i \leq c_T(l+1)^d \cdot 2^d\) and (2.56).

Based on the factorization formula (2.62), we derive now the multiplication algorithm of the matrix \(\hat{S}^L\) by a vector \(x\).

```
input : blockwise storage of \(S^L = (S^L_{\ell'})_{0 \leq \ell', \ell \leq L}\) (sparse), \(x = (x^L_{\ell})_{1 \leq \ell \leq \ell \leq L}\)
output: \(\hat{S}^L x\)

1 for \(\ell\) satisfying \(\sum_{\nu=1}^{k} l_\nu \leq L\) do
2 initialize \((\hat{S}^L x)^{\ell} := 0\)
3 for \(\ell'\) satisfying \(\sum_{\nu=1}^{k} l_\nu' \leq L\) do
4 compute \(y^\ell_{L} := T_{L,\ell',\ell}^L \cdot T_{L,\ell',\ell-1}^L \cdots T_{L,\ell',1}^L \cdot x^L_{\ell'}\)
5 update \((\hat{S}^L x)^{\ell} := (\hat{S}^L x)^{\ell} + y^\ell_{L}\)
6 end
7 end
```

\textbf{Algorithm 1: Sparse tensor matrix-vector multiplication}

\textbf{Remark 2.25.} The order in the product on line 5: of Algorithm 1 is essential for the efficiency of the multiplication.

\textbf{Remark 2.26.} Due to (2.60), (2.61), the multiplication of \(T_{L,\ell',\ell}^L\) by a vector \(x\) is done by multiplying the blocks of \(x\) by \(S^L_{\ell',\ell}\).

We estimate next the complexity of Algorithm 1.

\textbf{Theorem 2.27.} Algorithm 1 performs the matrix-vector multiplication \(x \mapsto \hat{S}^L x\) using at most \(O((\log N_L)^{kd+2k-2} N_L)\) floating point operations, and it requires only storage of the stiffness matrix \(S^L\) corresponding to the case \(k = 1\) (mean field problem) and of \(x\).

Proof. Due to (2.55), (2.62) we write
\[
(\hat{S}^L x)^{\ell} = \sum_{L<L',L} \sum_{\nu=1}^{k} S^L_{\ell',\ell} \cdot x^L_{\ell'} = \sum_{L<L',L} T_{L,\ell',\ell}^L \cdot T_{L,\ell',\ell-1}^L \cdots T_{L,\ell',1}^L \cdot x^L_{\ell'}.
\]

The multiplication under the summation can be performed using at most
\[
\#_{L,L'} := \sum_{q=1}^{k} \text{nnz}(T_{L,\ell',q}^L)
\]
(2.64)
2.4. ITERATIVE SOLUTION AND COMPLEXITY

floating point operations. From (2.63) we obtain that

\[ \#UL \leq c_{T,k,\psi} \sum_{q=1}^{k} (l_{\sigma(1)} + 1)^d \cdots (l_{\sigma(q-1)} + 1)^d \cdot (\min(l_{\sigma(q)}, l'_{\sigma(q)}) + 1)^{d-1}. \]

\[ \cdot (l'_{\sigma(q+1)} + 1)^d \cdots (l'_{\sigma(k)} + 1)^d \cdot 2^{d} \left( \sum_{\sigma=q}^{k-1} l_{\sigma(q)} + \max\{l_{\sigma(q)}, l'_{\sigma(q)}\} + \sum_{\sigma=q+1}^{k} l'_{\sigma(q)} \right). \]

Using the defining property (2.57) of \( \sigma = \sigma_{l'} \) we deduce that for \( L \geq 1 \),

\[ \#UL \leq c_{T,k,\psi} \left( \max\{\|l\|, \|l'\|\} \right)^{dk-1} \cdot 2^{d} \max\{\|l\|, \|l'\|\}. \]

The block \((\hat{S}^L x)_l\) can be computed using \( \sum_{l'} \#UL \) operations. Finally, the number of operations needed to perform \( x \to \hat{S}^L x \) (collect all blocks \((\hat{S}^L x)_l\) for all \( l \)) admits the upper estimate

\[ c_{T,k,\psi} \sum_{l=0}^{L} \sum_{l' \leq l} \left( \max\{\|l\|, \|l'\|\} \right)^{dk-1} \cdot 2^{d} \max\{\|l\|, \|l'\|\}. \]

Since for a given \( l \geq 0 \) the equation \( \|l\| = l \) has exactly \( \binom{l+k-1}{k-1} \leq (l+1)^{k-1} \) solutions \( l \in \mathbb{N}^k \), we conclude

\[ \#\text{flops}(x \to \hat{S}^L x) \leq c_{T,k,\psi} \sum_{l=0}^{L} (l+1)^{dk+2k-2} \cdot 2^{dl} \leq c_{T,k,\psi}(\log N_L)^{kd+2k-2} N_L. \]

But to Proposition 2.19, the number of steps required by the CG algorithm to compute the discrete solution up to a prescribed accuracy is bounded once we use the solution at level \( L-1 \) as initial guess of the solution at level \( L \). Thus it holds

**Theorem 2.28.** The deterministic problem (2.12) for the \( k \)-th order moment \( \mathcal{M}^k(u) \in H^{s+1}(D^k) \cap H_0^k(D^k) \) of the random solution \( u \) to (1.10) is solvable by a sparse tensor product FEM with \( N_L \) dof’s in \( D \) and for any \( L \geq k-1 \) at a cost of

\[ c_{T,k,\psi}(\log N_L)^{kd+2k-2} N_L \]  

for \( k \) floating point operations, with a memory

\[ c_{T,k}(\log N_L)^{k-1} N_L, \]

for a relative accuracy of

\[ c_{a,u,T,k}(\log N_L)^{(k-1)/2} N_L^{-\delta}, \]

where \( \delta = \min\{p, s\}/d \).

Note that, up to logarithmic terms, the estimates (2.65), (2.66), (2.67) are asymptotically as \( L \to \infty \) similar to those of the case \( k = 1 \) (mean field problem). Further, all constants involved depend exponentially on \( k \).
Chapter 3

Stochastic Coefficient

Under the uniform positivity Assumption 1.8 we investigate equation (1.10),

\[ \Delta_{a(\omega)} u(\omega) = f(\omega) \quad \text{in } H^{-1}(D) \quad P\text{-a.e. } \omega \in \Omega \]

using a perturbation argument. We describe the stochastic diffusion coefficient \( a \) as a random fluctuation \( r \) around a deterministic expectation (mean field) \( e \) and separate the deterministic and stochastic parts in \( r \) by a tensorial representation. Under regularity assumptions on \( r(x,\omega) \) in the physical variable \( x \) (but not on the size of \( r \), apart from the standard one ensuring the well-posedness of (1.10)) we prove that arbitrary (weighted) moments of the solution \( u \) to (1.10) can be computed in essentially the same complexity as the solution of one deterministic diffusion problem in physical domain \( D \). Besides, the resulting algorithm requires only a deterministic solver for diffusion problems in \( D \) and it is partially parallelizable.

3.1 Random Fields

We consider a splitting of the diffusion coefficient into a deterministic expectation \( e \) and a random fluctuation \( r \),

\[ a(x,\omega) = e(x) + r(x,\omega) \quad \forall (x,\omega) \in D \times \Omega \]  

with a positive \( e \in L^\infty(D) \) (not necessarily equal to \( \int a(\cdot,\omega) dP(\omega) \)),

\[ 0 < e_- \leq e(x) \leq e_+ < \infty \quad \forall x \in D. \]

It follows that the random fluctuation \( r \in L^\infty(D \times \Omega) \) too, and we require that the fluctuation be pointwise smaller than the expectation.

Assumption 3.1. For the representation (3.1) it holds

\[ 0 \leq \nu := \text{esssup}_{x \in D} \frac{\|r(x,\cdot)\|_{L^\infty(\Omega)}}{e(x)} < 1. \]  

(3.3)
Remark 3.2. The constant expectation choice

\[ e(x) := (a_- + a_+)/2 \quad \forall x \in D \]  

satisfies Assumption 3.1 with \( \nu \leq (a_+ - a_-)/(a_+ + a_-) < 1 \).

The more natural (from a statistical point of view) choice

\[ e(x) := \int a(x, \omega) dP(\omega) \quad \forall x \in D \]  

satisfies (3.3) if the density function of \( r(x, \cdot) \) is even for any \( x \in D \) (positive and negative fluctuations occur with equal probabilities).

Concerning the fluctuation term \( r \) we formulate also a modelling assumption as well as a condition of regularity in the physical variable.

Assumption 3.3. The random fluctuation \( r \) can be represented in \( L^\infty(D \times \Omega) \) as a convergent series

\[ r = \sum_{m=1}^{\infty} \phi_m \otimes X_m \]  

with known \( \phi_m \in C^\infty(D) \) (deterministic), \( X_m \in L^\infty(\Omega) \) (stochastic).

The representation formula (3.6) describes the tensor product nature of the random field \( r \), satisfying therefore the separation ansatz in the deterministic and stochastic variables \( x \in D \) and \( \omega \in \Omega \) respectively. A standard representation of this type is the Karhunen-Loève expansion (see e.g. [Loève77], [Loève78]), and a further example is given below (see Example 3.6).

The regularity of the random field \( r \) is quantified by the convergence rate of the series (3.6).

Assumption 3.4. The random fluctuation \( r \) admits a representation (3.6) for which there exist constants \( c_r, c_{1,r}, \alpha > 0 \) such that

\[ \| \phi_m \otimes X_m \|_{L^\infty(D \times \Omega)} \leq c_r \exp(-c_{1,r}m^\alpha) \quad \forall m \in \mathbb{N}_+. \]  

The regularity Assumption 3.4 holds for instance for random fields \( r \) which are analytic in the physical variable, with values in \( L^\infty(\Omega) \). This follows from standard approximation theory of analytic functions (see e.g. [Dav63]).

Definition 3.5. For an arbitrary Banach space \( B \) and an open set \( U \subseteq \mathbb{C}^n \) we denote by \( \mathcal{A}(U, B) \) the space of \( B \)-valued complex analytic functions on \( U \) which are also continuous on the closure of \( U \). Note that \( \mathcal{A}(U, B) \) equipped with

\[ \| u \|_{\mathcal{A}(U, B)} := \sup_{x \in U} \| u(x) \|_B \quad \forall u \in \mathcal{A}(U, B) \]

becomes a Banach space.
Example 3.6. If \( D \subset [-1, 1]^d \) and \( r \in \mathcal{A}([-1, 1]^d, L^\infty(\Omega)) \), then a representation (3.6) exists with \((\phi_m)_{m \in \mathbb{N}_+}\) the Legendre polynomials in \([-1, 1]^d\) (tensor products of standard Legendre polynomials in \([-1, 1]\), scaled to have \(L^2\)-norm equal to 1) and
\[
X_m(\omega) := \int_{[-1, 1]^d} r(x, \omega)\phi_m(x) \, dx \quad P\text{-a.e. } \omega \in \Omega, \forall m \in \mathbb{N}_+.
\] (3.8)
Moreover, (3.7) holds with \( \alpha = 1/d \) and \( c_{1,r} \) depending on the size of the analyticity domain of \( r \) in a complex neighbourhood of \([-1, 1]^d\).

Remark 3.7. By knowing \( X_m \) for \( m \in \mathbb{N}_+ \) we understand that any mixed moment of any finite subset of the random variable family \((X_m)_{m \in \mathbb{N}_+}\) is (computationally) available. For instance, for the construction presented in Example 3.6, the analyticity assumption on \( r \) ensures that the integral on the r.h.s. of (3.8) can be estimated using a quadrature rule \((x_q, w_q)_{q \in Q}\) in \([-1, 1]^d\). It follows
\[
X_m(\cdot) \preceq \sum_{q \in Q} r(x_q, \cdot)\phi_m(x_q)w_q. \tag{3.9}
\]

Given the realisations \((a(\cdot, \omega))_{\omega \in \Omega_0}\), an arbitrary \( k \)-th order mixed moment of the family \((X_m)_{m \in \mathbb{N}_+}\) is then estimated by the mean
\[
\int_{\Omega} X_{m_1}(\omega)X_{m_2}(\omega) \cdots X_{m_k}(\omega) \, dP(\omega) \succeq \frac{1}{|\Omega_0|} \sum_{\omega \in \Omega_0} \prod_{j=1}^k \left( \sum_{q \in Q} r(x_q, \omega)\phi_{m_j}(x_q)w_q \right). \tag{3.10}
\]

Remark 3.8. The following arguments, proofs and algorithms carry over to the case of a random fluctuation \( r \) satisfying Assumption 3.3 only locally on \( D \) (i.e. if a representation formula (3.6) holds locally on a finite covering of \( D \)). Moreover, a simple scaling argument shows that we can assume w.l.o.g. from now on that \( D \subset [-1, 1]^d \).

Since computations can handle only finite data sets, we truncate the infinite expansion (3.6) of \( r \) and introduce for any \( M \in \mathbb{N} \cup \{\infty\} \) (convention: \( r_0 = 0 \)) the \( M \)-term truncated fluctuation
\[
r_M := \sum_{m=1}^M \phi_m \otimes X_m, \quad a_M := e + r_M. \tag{3.11}
\]
The error introduced thereby in the solution of (1.10) is then controlled due to the Strang Lemma.

Proposition 3.9. If Assumption 3.4 holds, then there exists \( M_r \in \mathbb{N} \) such that for any \( M \geq M_r \) the problem of finding \( u_M \in L^0(\Omega, H^1(D)) \) such that
\[
\Delta a_M(\omega)u_M(\omega) = f(\omega) \quad \text{in } H^{-1}(D) \quad P\text{-a.e. } \omega \in \Omega
\] (3.12)
is well-posed and
\[
\|u(\omega) - u_M(\omega)\|_{H^1_0(D)} \leq c_{e,r} \exp(-c_{1,r}M^\alpha)\|f(\omega)\|_{H^{-1}(D)} \quad P\text{-a.e. } \omega \in \Omega. \tag{3.13}
\]
Remark 3.10. If \((X_m)_{m \in \mathbb{N}_+}\) is an independent family of random variables with vanishing mean (e.g. for the (3.5) choice of \(e\)), then \(M_r\) can be chosen equal to 0, due to the fact that problem (3.12) becomes elliptic uniformly in \(M \in \mathbb{N}\). This can be easily seen by taking the conditional expectation of (1.15) w.r.t. the \(\sigma\)-algebra generated by \(X_1, X_2, \ldots, X_M\).

Throughout this section we will constantly make use of the mixed moments of vector valued random fields, which are defined as follows.

Definition 3.11. If \(H\) is a separable Hilbert space, \(j \in \mathbb{N}_+ \cup \{\infty\}\), \(p = (p_1, p_2, \ldots, p_j) \in L^k(\Omega, H)^j\) and \(n = (n_1, n_2, \ldots, n_j) \in \mathbb{N}^j\) with \(\|n\|_1 := n_1 + n_2 + \cdots + n_j = k\), then
\[
M^n(p) := \int_{\Omega} p_1(\omega) \otimes \cdots \otimes p_1(\omega) \otimes \cdots \otimes p_j(\omega) \otimes \cdots \otimes p_j(\omega) dP(\omega) \in H \otimes \cdots \otimes H \quad (3.14)
\]
is called the mixed moment of \(p\) of order \(n\).

Further notations will be \(1_j := (1, 1, \ldots, 1)\) and, for \(m \in \mathbb{N}_+^j\), \(X_m := (X_{m_1}, X_{m_2}, \ldots, X_{m_j})\).

### 3.2 Discretization

Considering the hierarchical FE space sequence \((V_L)_{L \in \mathbb{N}}\) given by (2.18) and satisfying Assumption 2.11, we formulate for any \(M \in \mathbb{N} \cup \{\infty\}\) and any FE discretization level \(L \in \mathbb{N}\) the discrete counterpart of (3.12),

\[
P_L : \text{Find } u_{M,L} \in L^0(\Omega, V_L) \text{ such that } \Delta_{a} u_{M,L}(\omega) = f(\omega) \text{ in } V_L^* P\text{-a.e. } \omega \in \Omega. \quad (3.15)
\]

From Assumption 3.4, \(P_L\) is well-posed for \(M \geq M_r\).

Assuming that the conditions ensuring a shift theorem at level \(s\) for the diffusion problem in \(D\) with coefficient \(a(\cdot, \omega)\) and data \(f(\cdot, \omega)\) are satisfied uniformly in \(\omega \in \Omega\), the discretization error satisfies the standard estimate.

Assumption 3.12. There exists \(s > 0\) such that \(a \in L^\infty(\Omega, C^s(\overline{D}))\) (following e.g. from \(e \in C^s(\overline{D})\) and \(r \in A(\overline{D}, L^\infty(\Omega))\), and \(\partial D \in C^{s+1}\), so that the deterministic diffusion problem in \(D\) with coefficient \(a\) satisfies a shift theorem at level \(s \geq 0\) (see, e.g. [GT01]).

Proposition 3.13. If \(f \in L^0(\Omega, H^{-1+s}(D))\) and Assumptions 2.11, 3.4, 3.12 hold, then
\[
\|u_M(\omega) - u_{M,L}(\omega)\|_{H^s(D)} \leq c_a s \|f(\omega)\|_{H^{-1+s}(D)} \Phi(N_L, s) \quad P\text{-a.e. } \omega \in \Omega, \quad (3.16)
\]
and for any \(M \geq M_r, L \in \mathbb{N}\). Here the functional \(\Phi\) quantifies the approximation property of the scale \((V_L)_{L \in \mathbb{N}}\) (see also Example 2.12).
3.3 Neumann Series Decomposition

Considering \( f \in L^0(\Omega, H^{-1}(D)) \) and the hierarchical FE space sequence (2.18), we formulate for any \( M \in \mathbb{N} \cup \{\infty\}, M \geq M_r \) and any FE discretization level \( L \in \mathbb{N} \) the following sequence of well-posed problems,

\[ P_L : \text{Find } u_{M,L} \in L^0(\Omega, V_L) \text{ such that} \]
\[ \Delta_a u_{M,L}(\omega) = f(\omega) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega, \quad (3.17) \]

\[ P_{0,L} : \text{Find } v_{0,L} \in L^0(\Omega, V_L) \text{ such that} \]
\[ \Delta_e v_{0,L}(\omega) = f(\omega) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega, \quad (3.18) \]

and for \( j \in \mathbb{N}_+ \)

\[ P_{j,L} : \text{Find } v_{j,L} \in L^0(\Omega, V_L) \text{ such that} \]
\[ \Delta_e v_{j,L}(\omega) = \text{div}(r_M(\omega) \nabla v_{j-1,L}(\omega)) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega. \quad (3.19) \]

Using basic \( H^1 \) estimates for the diffusion problem in \( D \) we show next the exponential decay of \( v_{j,L} \), the solution to (3.19), in the energy norm as \( j \to \infty \), and the exponential convergence of the series with general term \( v_{j,L} \) to \( u_{M,L} \) solution to (3.17). To this end, we introduce, for any \( \epsilon \in \mathbb{C} \),

\[ a_{M,\epsilon} := e + \epsilon r_M \in L^\infty(D \times \Omega) \quad (3.20) \]

and study the diffusion problem with stochastic coefficient \( a_{M,\epsilon} \).

**Proposition 3.14.** Under Assumptions 3.1, 3.4, there exists \( \delta_\nu > 0 \) such that the problem of finding \( u_{M,L,\epsilon} \in L^0(\Omega, V_L \otimes \mathbb{C}) \) with

\[ -\Delta_{a_{M,\epsilon}} u_{M,L,\epsilon}(\omega) = f(\omega) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega, \quad (3.21) \]

is well-posed uniformly in \( M \geq M_r, L \in \mathbb{N} \) and \( \epsilon \in S \) where \( S \) denotes the complex strip

\[ S := [-1 - \delta_\nu, 1 + \delta_\nu] \times i\mathbb{R} \subset \mathbb{C}. \quad (3.22) \]

**Proof.** For notational ease we drop \( \omega \) from all notations. The real and imaginary parts \( \Re u_{M,L,\epsilon} \in V_L \) and \( \Im u_{M,L,\epsilon} \in V_L \) should solve \( P\text{-a.e. } \omega \in \Omega \) the system equivalent to (3.21),

\[
\begin{cases}
\Delta_{\Re a_{M,\epsilon}} \Re u_{M,L,\epsilon} - \Delta_{\Im a_{M,\epsilon}} \Im u_{M,L,\epsilon} = f \quad \text{in } V_L^* \\
\Delta_{\Im a_{M,\epsilon}} \Re u_{M,L,\epsilon} + \Delta_{\Re a_{M,\epsilon}} \Im u_{M,L,\epsilon} = 0 \quad \text{in } V_L^* 
\end{cases}
\quad (3.23)
\]

Choosing an ONB in \( V_L \) (equipped with the \( H^1_0(D) \) norm), (3.23) becomes a linear system whose matrix \( A \) has the form

\[ A = \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix} \quad (3.24) \]
with $A_1 \in \mathbb{R}^{N_L \times N_L}$ the real-valued symmetric matrix associated to the bilinear form of $\Delta_{a_{M,e}}$ in $V_L$, and $A_2 \in \mathbb{R}^{N_L \times N_L}$ the real-valued symmetric matrix associated to the bilinear form of $\Delta_{a_{M,e}}$ in $V_L$. Noting that

$$R_{a_{M,e}} = e + (\Re)e_r$$

is bounded away from zero for $e$ in $S$ given by (3.22) with $\delta_r$ small enough we deduce that $A_1$ has (positive) spectrum, bounded away from 0 uniformly in $\omega \in \Omega$, $L \in \mathbb{N}$ and $M \geq M_r$. Since

$$\left( \begin{array}{c} \alpha \\ \beta \end{array} \right) \left( \begin{array}{c} \alpha \\ \beta \end{array} \right) = \langle A_1 \alpha, \alpha \rangle_{\mathbb{R}^{N_L}} + \langle A_1 \beta, \beta \rangle_{\mathbb{R}^{N_L}}$$

we obtain that the spectrum of $A$ is bounded away from 0 uniformly in $\omega \in \Omega$, $L \in \mathbb{N}$ and $M \geq M_r$.

**Proposition 3.15.** Under Assumptions 3.1, 3.4, the mapping

$$S \ni e \rightarrow u_{M,L,e}(\omega) \in H^1_0(D)$$

with $u_{M,L,e}$ solution to (3.21) is holomorphic $P$-a.e. in $\Omega$ for any $M \geq M_r$, $L \in \mathbb{N}$, with $M_r$ as in Proposition 3.9.

**Proof.** Fixing $e_0 \in S$, we show that $u_{M,L,e}$ is holomorphic as a function of $e$ in a neighbourhood of $e_0$ by explicitly constructing the solution of (3.21) as a Taylor expansion around $e_0$. To this end we consider the following sequence of well-posed problems

$$\Delta_{a_{M,e_0}}(\omega)u_{M,L,e_0}(\omega) = f(\omega) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega,$$

and for $j \in \mathbb{N}_+$

$$\Delta_{a_{M,e_0}}(\omega)v_{j,L,e_0}(\omega) = \text{div}(r_M(\omega)\nabla v_{j-1,L,e_0}(\omega)) \quad \text{in } V_L^* \quad P\text{-a.e. } \omega \in \Omega,$$

where by convention $v_{0,L,e_0} := u_{M,L,e_0}$. From (3.29) with $v_{j,L,e}$ as test function we deduce that, $P$-a.e. in $\Omega$ it holds,

$$\|\nabla v_{j,L,e_0}(\omega)\|_{L^2(\Omega)} \leq \frac{\|r_M\|_{L^\infty(D \times \Omega)}}{c_{e_0}} \|\nabla v_{j-1,L,e_0}(\omega)\|_{L^2(\Omega)} \quad \forall j \in \mathbb{N}_+.$$  

Note that the existence of $c_{e_0} > 0$ is ensured by Proposition 3.14. From (3.30) we obtain the exponential decay of $|\rho|^j\|v_{j,L,e_0}(\omega)\|_{H_0^1(D)}$ with $j \to \infty$, uniformly $P$-a.e. in $\Omega$ for $\rho$ in a small complex neighbourhood of 0 depending on $e_0$ and $\|r\|_{L^\infty(D \times \Omega)}$. In particular the series

$$\sum_{j \in \mathbb{N}} \rho^j v_{j,L,e_0}(\omega)$$

converges in $H^1_0(D)$, uniformly $P$-a.e. in $\Omega$ for $\rho$ in a small complex neighbourhood of 0. Moreover, the sum solves the diffusion problem with stochastic coefficient $a_{M,e_0+\rho}$, as one can easily see by summing (3.28) and (3.29) for $j \in \mathbb{N}_+$. The value of (3.31) is therefore equal to $u_{M,L,e_0+\rho}(\omega)$ $P$-a.e. in $\Omega$, which concludes the proof. □
3.4. DISCRETE MIXED MOMENT PROBLEM

The domain of analyticity of the mapping (3.27) contains therefore the closed disc of center 0 and radius 1 in the complex plane, which ensures that the distance between \( \epsilon_0 = 0 \) and the singularities of (3.27) is strictly larger than 1. As a consequence, the Taylor expansion of (3.27) around \( \epsilon_0 = 0 \) converges exponentially in \( \rho = 1 \).

**Proposition 3.16.** If \( f \in L^0(\Omega, H^{-1+s}(D)) \) and Assumptions 3.1, 3.4 hold, there exists \( c_{1,a} > 0 \) such that

\[
\|v_{j,L}(\omega)\|_{H^s_0(D)} \leq c_a \|f(\omega)\|_{H^{-1}(D)} \exp(-c_{1,a}j) \quad P\text{-a.e. } \omega \in \Omega, \tag{3.32}
\]

and

\[
\|u_{M,L}(\omega) - \sum_{j=0}^{n} v_{j,L}(\omega)\|_{H^s_0(D)} \leq c_a \|f(\omega)\|_{H^{-1}(D)} \exp(-c_{1,a}n) \quad P\text{-a.e. } \omega \in \Omega, \tag{3.33}
\]

and for any \( M \geq M_r, j, n, L \in \mathbb{N} \).

**Proof.** Taylor expansion of (3.27) around \( \epsilon_0 = 0 \) gives,

\[
u_{j,L,0}(\omega) = \sum_{j \in \mathbb{N}} \rho^j v_{j,0}(\omega) \tag{3.34}
\]

so that

\[
v_{j,L,0}(\omega) = \frac{1}{j!} \partial^j u_{M,L,\epsilon}(\omega)|_{\epsilon=0}. \tag{3.35}
\]

Noting that \( v_{j,L,0} = v_{j,L} \) for any \( j \in \mathbb{N}_+ \) and \( u_{M,L,1} = u_{M,L} \), all we have to check is the exponential convergence in \( H^s_0(D) \) of the series on the l.h.s. of (3.34) for \( \rho = 1 \). But this is equivalent to the exponential decay of \( v_{j,L,0} \) as \( j \to \infty \) in the \( H^s_0(D) \) norm, which in turn follows from (3.35) and the Cauchy representation formula

\[
\partial^j u_{M,L,\epsilon}(\omega)|_{\epsilon=0} = \frac{j!}{2\pi i} \int_{\Gamma} \frac{u_{M,L,z}(\omega)}{z^{j+1}} \, dz \quad \forall j \in \mathbb{N}
\]

applied on an arbitrary circle \( \Gamma \) of center 0 and radius strictly larger than 1, contained in the strip \( S \).

\[\square\]

### 3.4 Discrete Mixed Moment Problem

Since our primary aim is the development and the analysis of an algorithm based on the Neumann expansion (3.18), (3.19) for the computation of the simplest (first/second order) statistics of \( u \) solution to (1.10), let us collect now in the next result the errors introduced by the discretization steps of the previous sections (modelling, FE discretization and Neumann truncation errors, see (3.13), (3.16) and (3.33)).
Proposition 3.17. If \( f \in L^1(\Omega, H^{-1+s}(D)) \) and Assumptions 3.1, 3.4, 3.12 hold, then

\[
\|u(\omega) - \sum_{j=0}^{n} v_{j,L}(\omega)\|_{H^s(D)} \leq 
\]

\[
\leq c_{a,s}\|f(\omega)\|_{H^{-1+s}(D)}(\Phi(N_L,s) + \exp(-c_{1,\alpha}M^\alpha) + \exp(-c_{1,\alpha}n)) \quad P\text{-a.e. } \omega \in \Omega, \tag{3.36}
\]

for any \( M \geq M_r, n, L \in \mathbb{N} \).

Note that, after averaging (3.36) over \( \omega \in \Omega \) and balancing the truncation orders \( M, n \) and the discretization level \( L \), the computation of the mean fields \( \mathcal{M}^1(v_{j,L}) \) for \( 0 \leq j \leq n \) produces also an approximation of \( \mathcal{M}^1(u) \). However, computing the moments \( \mathcal{M}^1(v_{j,L}) \) for \( 0 \leq j \leq n \) is not an obvious task. Under additional regularity assumptions on the random fluctuation \( r \), we develop next an algorithm which makes this computation possible in nearly optimal complexity (compare Definition 1.1). It turns out that this computation has to be based on knowledge of higher order mixed moments of the pair \( (r_M, f) \). Assuming these are available, we derive a general algorithm for the computation of the mixed moments associated to the pair \( (r_M, u) \) (and not only of \( \mathcal{M}^1(u) \)). This algorithm will be later used in the context of second (and higher) order moment computation.

Assumption 3.18. There exists an open complex neighbourhood \( U \subset \mathbb{C}^d \) of \( \overline{D} \) such that Assumption 3.4 is satisfied and the decay estimate (3.7) holds in \( L^\infty(U \times \Omega) \) (in particular, \( \phi_m \in \mathcal{A}(U) \) for all \( m \in \mathbb{N}_+ \)).

Remark 3.19. Assumption 3.18 implies \( r \in \mathcal{A}(U, L^\infty(\Omega)) \) and the convergence of the series (3.6) in \( \mathcal{A}(U, L^\infty(\Omega)) \). Reciprocally, if \( r \in \mathcal{A}(U', L^\infty(\Omega)) \) with \( U \subset U' \subset \mathbb{C}^d \) (which holds e.g. for a diffusion coefficient \( a \) with Gaussian covariance), then Assumption 3.18 is satisfied with (3.6) the Legendre/Karhunen-Loève expansion of \( r \) (compare Example 3.6).

Definition 3.20. If \( X \) is a bounded linear operator between two Banach spaces \( B_1 \) and \( B_2 \), and \( U \subset \mathbb{C}^n \) is an open set, let \( \mathcal{A}(U, B_i) \) be the Banach space of \( B_i \)-valued functions which are continuous on \( U \) and analytic in \( U \) (i = 1, 2). We define \( \text{Id}_U \otimes X \) to be the bounded linear operator given by

\[
\text{Id}_U \otimes X : \mathcal{A}(U, B_1) \to \mathcal{A}(U, B_2), \quad ((\text{Id}_U \otimes X)h)(z) := X(h(z)) \quad \forall z \in \overline{U}.
\]

Remark 3.21. The direct sum of \( d \) copies of the trace operator \( \Xi \) defined in Lemma 5.1 induces by restriction to functional spaces on \( U \) a bounded linear operator denoted for notational ease again by \( \Xi \),

\[
\Xi : \mathcal{A}(U, L^2(D)^d) \to L^2(D)^d, \quad (\Xi(u))(x) = u(x, x) \quad \lambda\text{-a.e. } x \in D. \tag{3.37}
\]

We formulate next for \( j \in \mathbb{N}_+ \) arbitrary the problem to which the mixed moment of \( (r_M, v_{j,L}) \) of order \( (j, 1) \) is solution. To this end we introduce the notation

\[
\Delta_{e,L}^{-1} : H^{-1}(D) \to V_L \tag{3.38}
\]
for the bounded FE solution operator of the homogeneous diffusion problem with coefficient \( e \) corresponding to the FE space \( V_L \) equipped with the \( H^0(D) \) norm.

Further, for any \( U \subseteq \mathbb{C}^d \) and any \( j \in \mathbb{N}_+ \) we define the product domain

\[
U^j := U \times U \times \cdots \times U \subset \mathbb{C}^d.
\]

**Definition 3.22.** For any \( j \in \mathbb{N}_+ \) and \( L \in \mathbb{N} \) we define the operator \( T_{j,L} \) as the following composition of linear operators,

\[
A(U^j, V_L) \xrightarrow{\text{Id}_{U^j} \otimes \nabla} A(U^j, L^2(D)^d) \xrightarrow{\text{Id}_{U^j-1} \otimes \epsilon} A(U^j-1, L^2(D)^d) \xrightarrow{\text{div}} A(U^j-1, H^{-1}(D)) \xrightarrow{\text{Id}_{U^j-1} \otimes \Delta_{e,L}^{-1}} A(U^j-1, V_L).
\]

Similarly, for any \( j \in \mathbb{N} \) and \( L \in \mathbb{N} \) we define the operator \( S_{j,L} \) given by

\[
A(U^j, H^{-1}(D)) \xrightarrow{\text{Id}_{U^j} \otimes \Delta_{e,L}^{-1}} A(U^j, V_L).
\]

It then holds

**Proposition 3.23.** The operator \( T_{j,L} \) defined by (3.39) is bounded uniformly in \( j, L \) and maps \( M^{(j,1)}(r_M, \nu_{j',L}) \) into \( M^{(j-1,1)}(r_M, \nu_{j'+1,L}) \) for any \( j' \in \mathbb{N} \).

Similarly, the operator \( S_{j,L} \) defined by (3.40) is bounded uniformly in \( j, L \) and maps \( M^{(j,1)}(r_M, f) \) into \( M^{(j,1)}(r_M, \nu_{0,L}) \).

**Proof.** Note first that the mixed moments have been introduced in Definition 3.11. Observe next that the second operator in (3.39) is bounded uniformly in \( j \) due to Lemma 5.1 (see Appendix) and Definition 3.20, while the others are trivially continuous, uniformly in \( L \).

The mapping property follows by taking the tensor product of \( P_{j'+1,L} \) (see (3.19)) with \( r_M(\omega) \otimes \cdots \otimes r_M(\omega) \) \((j-1)\) times and averaging the resulting equation over \( \omega \in \Omega \).

The mapping property of \( S_{j,L} \) follows then analogously, by taking the tensor product of \( P_{0,L} \) (see (3.18)) with \( r_M(\omega) \otimes \cdots \otimes r_M(\omega) \) \((j-1)\) times and averaging the resulting equation over \( \omega \in \Omega \).

Proposition 3.23 shows in particular that the first order moment (expectation) \( M^{(1)}(\nu_{j,L}) = M^{(0,1)}(r_M, \nu_{j,L}) \) follows by successive composition of operators of type (3.39) from the higher order moment \( M^{(j,1)}(r_M, \nu_{0,L}) \). This in turn is obtained by mapping \( M^{(j,1)}(r_M, f) \) via \( S_{j,L} \). The discrete problem solved by \( M^{(1)}(\nu_{j,L}) \) is therefore given by

**Corollary 3.24.** For any \( j, j' \in \mathbb{N} \) it holds

\[
M^{(j,1)}(r_M, \nu_{j,L}) = (T_{j'+1,L} \circ T_{j'+2,L} \circ \cdots \circ T_{j'+j,L} \circ S_{j'+j,L})(M^{(j'-1,1)}(r_M, f)).
\]

In particular, for \( j' = 0 \),

\[
M^{(1)}(\nu_{j,L}) = (T_{1,L} \circ T_{2,L} \circ \cdots \circ T_{j,L} \circ S_{j,L})(M^{(1,1)}(r_M, f)).
\]
Note that (3.41) can be viewed as a recursion of parametric deterministic diffusion problems in \( D \) where after each recursion step taking the trace via \( \Xi \) reduces the dimension of the parameter space by \( d \).

### 3.4.1 Mixed Moment Algorithm

From Corollary 3.24 and Proposition 3.17 it follows that the computation of the exact mean field \( \mathcal{M}^1(u) \) requires knowledge of the mixed moments \( \mathcal{M}^{(j,1)}(r_M, f) \) for \( j \in \mathbb{N} \). The more such moments are available, the higher the truncation order of the Neumann series and the better the expected approximation of \( \mathcal{M}^1(u) \).

Concerning the mixed moments \( \mathcal{M}^{(j,1)}(r_M, f) \) we have the following representation formula, and estimates following from Assumption 3.18.

**Proposition 3.25.** If \( j \in \mathbb{N} \) and \( r \) satisfies Assumption 3.18, then for

\[
\mathcal{M}^{(j,1)}(r_M, f) = \int_{\Xi} r_M(\omega) \otimes \cdots \otimes r_M(\omega) \otimes f(\omega) \, dP(\omega) \in \mathcal{A}(U^J, H^{-1}(D))
\]

it holds

\[
\mathcal{M}^{(j,1)}(r_M, f) = \sum_{m \in \mathbb{N}_+^j, \|m\|_\infty \leq M} \mathcal{M}^{j+1}(X_m, f) \otimes \phi_m,
\]

where for any index \( m = (m_1, m_2, \ldots, m_j) \in \mathbb{N}_+^j, \|m\|_\infty := \max_{1 \leq k \leq j} |m_k|, \)

\[
\phi_m := \phi_{m_1} \otimes \phi_{m_2} \otimes \cdots \otimes \phi_{m_j} \in \mathcal{A}(U^J),
\]

and (in view of Definition 3.11)

\[
\mathcal{M}^{j+1}(X_m, f) = \int_{\Xi} X_{m_1}(\omega) X_{m_2}(\omega) \cdots X_{m_j}(\omega) f(\omega) \, dP(\omega) \in H^{-1}(D).
\]

Further, the general term of the expansion (3.44) satisfies

\[
\|\mathcal{M}^{j+1}(X_m, f) \otimes \phi_m\|_{\mathcal{A}(U^J, H^{-1}(D))} \leq c_f \exp(-c_{1,r} \|m\|_\alpha^\alpha) \|f\|_{L^1(\Omega, H^{-1}(D))}
\]

with \( \|m\|_\alpha := (m_1^\alpha + m_2^\alpha + \cdots + m_j^\alpha)^{1/\alpha} \).

Equation (3.44) and the definitions of all operators involved in (3.42) enable us to derive the following algorithms for the computation of the mixed moments of the pairs \((r_M, u_{M,L})\) and \((r_M, u_{M,L})\) respectively, once a solver for the deterministic diffusion problem in \( D \) is available.
3.4. DISCRETE MIXED MOMENT PROBLEM

input: M, j, j', L
- approximation of $\mathcal{M}^{(j'+j,1)}(r_M, f)$ of the form

$$\mathcal{M}_{\text{app}}^{(j'+j,1)}(r_M, f) = \sum_{m \in \mathbb{N}_+^{j'+j}, \|m\|_{\infty} \leq M} h_m \phi_m \quad \text{with } h_m \in H^{-1}(D)$$

% e.g. $h_m := \mathcal{M}^{1,j+1}(X_m, f)$ for exact representation of $\mathcal{M}^{(j'+j,1)}(r_M, f)$
- solver of diffusion problem in $D$ with deterministic coefficient $e$ and FE space $V_L$

output: approximation of $\mathcal{M}^{(j,1)}(r_M, v_{j,L})$ of the form

$$\mathcal{M}_{\text{app}}^{(j,1)}(r_M, v_{j,L}) = \sum_{m \in \mathbb{N}_+^{j'}, \|m\|_{\infty} \leq M} g_m \phi_m$$

with $g_m \in V_L$

1. compute for all $m \in \mathbb{N}_+^{j'+j}, \|m\|_{\infty} \leq M$

$$g_m := \Delta_{e,L}^{-1} h_m \in V_L$$

2. for $k = j' + j : -1 : j' + 1$ do

3. compute for all $m \in \mathbb{N}_+^{k-1}, \|m\|_{\infty} \leq M$

$$g_m := \Delta_{e,L}^{-1} \left( \text{div} \sum_{m=1}^{M} \phi_m \nabla g_{(m,m)} \right) \in V_L$$

4. end

5. compute and return

$$\mathcal{M}_{\text{app}}^{(j',1)}(r_M, v_{j,L}) := \sum_{m \in \mathbb{N}_+^{j'}, \|m\|_{\infty} \leq M} g_m \phi_m$$

Algorithm 2: $\mathcal{M}_{\text{app}}^{(j',1)}(r_M, v_{j,L}) = \text{Algorithm 2} \left[ M, j, j', L, \mathcal{M}_{\text{app}}^{(j'+j,1)}(r_M, f) \right]$
CHAPTER 3. STOCHASTIC COEFFICIENT

| input: | • \(M, n, j', L\)  
| • approximations of \(\mathcal{M}^{(j'+j,1)}(r_M, f)\) for all \(0 \leq j \leq n\), of the form  
| \[ \mathcal{M}^{(j',1)}_{\text{app}}(r_M, f) = \sum_{\substack{m \in \mathbb{N}^j \cap \mathbb{N}^j \cap j \leq M}} h_m \phi_m \quad \text{with} \ h_m \in H^{-1}(D) \]  
| % e.g. \(h_m := \mathcal{M}^{(j'+j+1)}(X_m, f)\) for exact representation of \(\mathcal{M}^{(j'+j,1)}(r_M, f)\)  
| • solver of diffusion problem in \(D\) with deterministic coefficient \(\varepsilon\) and \(\text{FE space } V_L\)  
| output: | • approximation of \(\mathcal{M}^{(j',1)}(r_M, u_L)\) of the form  
| \[ \mathcal{M}^{(j',1)}_{\text{app}}(r_M, u_L) = \sum_{m \in \mathbb{N}^j \cap m \leq M} g_m \phi_m \]  
| with \(g_m \in V_L\)  

1. compute for all \(j = 0 : 1 : n\)  
\[ \mathcal{M}^{(j',1)}_{\text{app}}(r_M, v_L) := \text{Algorithm 2} \left[ M, j, j', L, \mathcal{M}^{(j'+j,1)}_{\text{app}}(r_M, f) \right] \]  

2. compute and return  
\[ \mathcal{M}^{(j',1)}_{\text{app}}(r_M, u_L) := \sum_{j=0}^{n} \mathcal{M}^{(j',1)}_{\text{app}}(r_M, v_L) \]

Algorithm 3: \(\mathcal{M}^{(j',1)}_{\text{app},n}(r_M, u_L) = \text{Algorithm 3} \left[ M, n, j', L, (\mathcal{M}^{(j'+j,1)}_{\text{app}}(r_M, f))_{0 \leq j \leq n} \right] \)

Proposition 3.26. For Algorithm 2 it holds  
\[ \mathcal{M}^{(j',1)}_{\text{app}}(r_M, v_L) = (T_{j'+1,L} \circ T_{j'+2,L} \circ \cdots \circ T_{j'+j,L} \circ S_{j'+j,L})(\mathcal{M}^{(j'+j,1)}_{\text{app}}(r_M, f)), \quad (3.47) \]

so that  
\[ \|\mathcal{M}^{(j',1)}_{\text{app}}(r_M, v_L) - \mathcal{M}^{(j',1)}_{\text{app}}(r_M, v_L)\|_{A(U^{j',1},H^1(D))} \leq \]  
\[ \leq c_{\text{al}}^j \|\mathcal{M}^{(j'+j,1)}_{\text{app}}(r_M, f) - \mathcal{M}^{(j'+j,1)}_{\text{app}}(r_M, f)\|_{A(U^{j'+j,1},H^{-1}(D))}. \quad (3.48) \]

Proof. Estimate (3.48) is a direct consequence of (3.47) and Proposition 3.23, so that it suffices to prove (3.47). To this end, we simply note that, in view of (3.40), applying \(S_{j'+j,L}\) to \(\mathcal{M}^{(j'+j,1)}_{\text{app}}(r_M, f)\) corresponds to step 1: of the algorithm, whereas (3.39) ensures that step 3: consists of successive applications of the operators \(T_{j'+j,L}, T_{j'+j-1,L}, \ldots, T_{j'+1,L}\).  

\[ \square \]
3.4. DISCRETE MIXED MOMENT PROBLEM

Proposition 3.27. For Algorithm 3 it holds

\[ \|M'(r_M, u_M, L) - M_{app,1}(r_M, u_M, L)\|_{A(U', H_0^1(D))} \leq \]
\[ \leq c_a \|f\|_{L^2(U, H^{-1}(D))} \|r_M\|_{L^\infty(U \times \Omega)} \exp(-c_1a_1) + \]
\[ + \sum_{j=0}^{n} c_n j ! \|M'(r_M, f) - M_{app,1}(r_M, f)\|_{A(U', H^{-1}(D))}. \quad (3.49) \]

Proof. The l.h.s. of (3.49) can be estimated from above using the triangle inequality by

\[ \|M'(r_M, u_M, L - \sum_{j=0}^{n} v_j, L)\| + \sum_{j=0}^{n} \|M'(r_M, v_j, L) - M_{app,1}(r_M, v_j, L)\| \quad (3.50) \]

(all norms in \( A(U', H_0^1(D)) \)). The conclusion follows then from the error estimates in Propositions 3.16, 3.26.

One important property of the mixed moments which can be immediately used to reduce the complexity of Algorithms 2, 3 is their symmetry w.r.t. \( m \). To describe it, we define on \( N_+^j \) the equivalence relation

\[ m \sim m' \iff \exists \sigma \text{ permutation of } \{1, 2, \ldots, j\} \text{ s.t. } m'_k = m_{\sigma(k)} \forall 1 \leq k \leq j, \quad (3.51) \]

so that for any \( m \in N_+^j \) the orbit \( O(m) \) of \( m \) is given by

\[ O(m) := \{ m' \in N_+^j | \exists \sigma \text{ permutation of } \{1, 2, \ldots, j\} \text{ s.t. } m'_k = m_{\sigma(k)} \forall 1 \leq k \leq j \}. \]

Due to Definition 3.11 it then clearly holds

Remark 3.28. The mixed moment \( M^{j+1}(X_m, f) \) is symmetric in \( m \), that is,

\[ m \sim m' \implies M^{j+1}(X_m, f) = M^{j+1}(X_{m'}, f) \in H^{-1}(D). \quad (3.52) \]

Remark 3.28 shows that reasonable approximations of the mixed moments of the pair \((r_M, f)\) used as input for Algorithm 2 should also satisfy the symmetry condition w.r.t. \( m \),

\[ m \sim m' \implies h_m = h_{m'} \in H^{-1}(D). \quad (3.53) \]

A simple backward induction argument on \( k \) shows then that this symmetry is preserved by step 1: and the recursion in step 3: of Algorithm 2.

Proposition 3.29. If the input data \((h_m)_{m \in N_+^j} \) is symmetric w.r.t. \( m \), then \( g_m \) given by steps 1: and 3: of Algorithm 2 is also symmetric in \( m \),

\[ m \sim m' \implies g_m = g_{m'} \in V_L \subset H_0^1(D). \quad (3.54) \]
Remark 3.30. The complexity of the mean field computation using Algorithm 3 with $j' = 0$ and exact input data $h_m = M^{j+1}(X_m, f)$ is in general superalgebraic. This follows by a simple comparison error estimate (3.36) vs. total number of deterministic problems which are to be solved in steps 1: and 3: of Algorithm 2. After choosing

$$M \sim L^{1/\alpha}, \quad n \sim L$$

in order to balance the contributions of the discretization steps to the total error estimate (3.36), Algorithm 3 requires solving (even after taking into account the symmetry of the $g_m$'s) at least

$$\frac{M^n}{n!} \geq N_L^{O((1/\alpha-1) \log N_L + 1)}$$

(3.56)
deterministic diffusion problems in $D$ at level $L$ (already in step 1; Algorithm 2 with $j = n$) for an accuracy of $O(N_L^{-\delta})$. Recalling that $\alpha = 1/d$ for analytic regularity of the perturbation $r$ (see also Example 3.6), (3.56) shows the superalgebraic complexity of Algorithm 3 in dimension $d \geq 2$. In view of the algebraic (quadratic) complexity of the Monte Carlo method, Algorithm 3 becomes then too expensive to use. Note however that a subexponential complexity upper bound can be easily shown for Algorithm 2 by counting arguments similar to those leading to (3.56).

The high complexity of Algorithm 3 can be significantly reduced by taking into account the analyticity assumption on $r$, through the estimate (3.46). We show that representation (3.44) can be sparsified, that is, many terms can be neglected (due to their small size as estimated by (3.46)), without affecting the total discretization error estimate (3.36). We consider thus for any $j \in \mathbb{N}$ an index set (the explicit construction will follow in section 3.4.2)

$$\mathcal{T}_j \subseteq \{ m \in \mathbb{N}^d_+ \mid \|m\|_\infty \leq M \}$$

(3.57)
(convention: $\mathbb{N}_0^d = \{0\}$, $\mathcal{T}_0 := \{0\}$) and apply Algorithms 2, 3 using as input the approximation of $M^{j'+j+1}(r_M, f)$ given by

$$M^{j'+j+1}_{\text{app}}(r_M, f) := \sum_{m \in \mathcal{T}_{j'+j}} h_m \otimes \phi_m,$$  

(3.58)
which amounts to simply setting the input coefficients $h_m := 0$ for $m \notin \mathcal{T}_{j'+j}$.

Defining

$$M^{(j)'(1)}_{\text{app}}(r_M, v_{j,l}) := (T_{j'+1,l} \circ T_{j'+2,l} \circ \cdots \circ T_{j'+j,l} \circ S_{j'+j,l})(M^{(j'+j+1)}_{\text{app}, T}(r_M, f)),$$

(3.59)
we rewrite Algorithms 2, 3 corresponding to the truncation index sets $(\mathcal{T}_{j})_{j \in \mathbb{N}}$ (see Algorithms 4, 5).

Concerning notations, we introduce for any $j, j' \in \mathbb{N}$ with $j \geq j'$ the projection operator onto the first $j'$ coordinates, $\Pi_{j,j'} : \mathbb{N}_+^d \rightarrow \mathbb{N}_+^{j'}$, and for any index set $\Gamma \subseteq \mathbb{N}_+^d$, we put

$$\Pi_{j,j'}(\Gamma) := \bigcup_{m \in \Gamma} \Pi_{j,j'}(m) = \{ m' \in \mathbb{N}_+^{j'} \mid \exists m \in \Gamma \text{ s.t. } m_k = m_k \text{ } \forall 1 \leq k \leq j' \} \subseteq \mathbb{N}_+^{j'}.$$
3.4. **DISCRETE MIXED MOMENT PROBLEM**

**Remark 3.31.** The approximation error estimates for Algorithms 4, 5 are the same as those given in Propositions 3.26 and 3.27, with approximate moments $\mathcal{M}_{\text{app}}$ replaced by sparsified approximate moments $\mathcal{M}_{\text{app},\mathcal{T}}$.

By sparsifying the moment representation using a convenient index family $(\mathcal{T}_j)_{j \in \mathbb{N}}$ we gain a substantial complexity reduction of Algorithm 3 (number of deterministic problems to be solved in step 1.; see also Remark 3.30), without deteriorating the error estimate (3.49). This will be shown in section 3.5, after the explicit construction of the index sets $(\mathcal{T}_j)_{j \in \mathbb{N}}$. Here we only observe that if the index sets $(\mathcal{T}_j)_{j \in \mathbb{N}}$ are compatible with the equivalence relation $\sim$, that is,

$$m \in \mathcal{T}_j \iff \mathcal{O}(m) \subseteq \mathcal{T}_j \quad \forall j \in \mathbb{N}, \quad (3.60)$$

and with the projection operators $\Pi$, that is,

$$\Pi_{j,j'}(\mathcal{T}_j) = \mathcal{T}_{j'} \quad \forall j, j' \in \mathbb{N}, j \geq j', \quad (3.61)$$

then the following complexity estimate for Algorithm 4 is a consequence of Proposition 3.29, which applies also in the case of sparsified input data.

**Proposition 3.32.** If the index family $(\mathcal{T}_j)_{j \in \mathbb{N}}$ satisfies the compatibility conditions (3.60), (3.61), then Algorithm 4 requires solving $|\mathcal{T}_{j'+j} / \sim | + |\mathcal{T}_{j'+j-1} / \sim | + |\mathcal{T}_{j'+j-2} / \sim | + \cdots + |\mathcal{T}_{j'} / \sim |$ deterministic diffusion problems in $D$ at FE discretization level $L$.

### 3.4.2 Sparsification

For a fixed $M \in \mathbb{N}$ and any $j, q \in \mathbb{N}$ we define the sparsification index sets

$$\mathcal{T}_{j,q} := \{ m \in \mathbb{N}_+ \mid \|m\|_\infty \leq M \text{ and } m \text{ has less than } q \text{ different entries} \}. \quad (3.62)$$

**Proposition 3.33.** Under Assumption 3.18 there exists $c_{2,r} > 0$ such that for any $j, q \in \mathbb{N}$ it holds

$$\| \sum_{m \in \mathbb{N}_+ \setminus \mathcal{T}_{j,q}}^{\mathcal{M}^{1_{j+1}}(X_m, f) \otimes \phi_m} \|_{L^\infty(U_j, H^{-1}(D))} \leq c_{2,r} \exp(-c_{2,r} q^{1+\alpha}) \| f \|_{L^1(\Omega, H^{-1}(D))}. \quad (3.63)$$

**Proof.** In view of (3.46) it suffices to estimate (convention: empty sum equals 0)

$$S_{j,q} := c_{2,r} \sum_{m \in \mathbb{N}_+ \setminus \mathcal{T}_{j,q}}^{\exp(-c_{1,r} \|m\|_\infty)} \| m \|_\infty.$$

To each $m \in \mathbb{N}_+^j$ with $\|m\|_\infty \leq M$ we associate the vector $n = (n_1, n_2, \ldots, n_M) \in \mathbb{N}^M$ defined by

$$\forall 1 \leq m \leq M \quad n_m := p \iff m \text{ has } p \text{ entries equal to } m. \quad (3.65)$$
Then $\|n\|_1 := n_1 + n_2 + \cdots + n_M = j$ and for any $n$ with $\|n\|_1 = j$ there are exactly $\frac{j!}{n_1!n_2! \cdots n_M!}$ indices $m \in \mathbb{N}_+^j$ with $\|m\|_\infty \leq M$ to which $n$ is associated via (3.65). Moreover, we have

$$m \in \Upsilon_{j,q} \iff n_m \geq 1 \text{ for less than } q \text{ different values of } m \text{ taken from } \{1, \ldots, M\}.$$  

(3.66)

With this notation, the sum $S_{j,q}$ in (3.64) can be written

$$S_{j,q} = c_r^j \sum_{\|n\|_1=j, \supp(n) \geq q} \frac{j!}{n_1!n_2! \cdots n_M!} \exp(-c_{1,r} \sum_{m=1}^M n_m m^a),$$

or, parametrizing the indices $n$ through their support and the values they assume there,

$$S_{j,q} = c_r^j \sum_{t=q}^{\min\{j,M\}} \sum_{1 \leq m_1 < \cdots < m_i \leq M} \sum_{n_{m_1}, \ldots, n_{m_i} \geq q} \frac{j!}{n_{m_1}! \cdots n_{m_i}!} \exp(-c_{1,r} \sum_{k=1}^i n_{m_k} m_k^a).$$  

(3.67)

The general term of the sum (3.67) can be estimated from above by

$$\frac{(j-i)!}{(n_{m_1} - 1)! \cdots (n_{m_i} - 1)!} \exp(-c_{1,r} \sum_{k=1}^i (n_{m_k} - 1)m_k^a) \frac{j!}{(j-i)!} \exp(-c_{1,r} \sum_{k=1}^i m_k^a).$$  

(3.68)

Using (3.68) in (3.67) and performing the innermost sum via the multinomial formula gives

$$S_{j,q} \leq c_r^j \sum_{t=q}^{\min\{j,M\}} \sum_{1 \leq m_1 < \cdots < m_i \leq M} \left( \sum_{k=1}^i \exp(-c_{1,r} m_k) \right)^{j-i} \frac{j!}{i!} \sum_{1 \leq m_1 < \cdots < m_i \leq M} \exp(-c_{1,r} \sum_{k=1}^i m_k^a).$$

(3.69)

Due to $m_k \geq k$ for any $1 \leq k \leq i$ we absorb the factorial in the exponential under the inner summation and obtain (with any $0 < c_{2,r} < c_{1,r}$)

$$S_{j,q} \leq c_r^j \sum_{t=q}^{\min\{j,M\}} c_{r}^{-i} \binom{j}{i} \sum_{1 \leq m_1 < \cdots < m_i \leq M} \exp(-c_{2,r} \sum_{k=1}^i m_k^a).$$

(3.70)

The conclusion follows then from Lemma 5.2 (see Appendix). □

Proposition 3.33 shows that the sparsification error in the mixed moment representation (3.44) with the index set $\Upsilon_{j,q}$ (3.62) is just as large as the largest neglected term, up to the constants involved.

A further obvious property of the index sets $(\Upsilon_{j,q})_{j \in \mathbb{N}}$ is their consistency with the equivalence relation $\sim$ and with the projection operators $\Pi$ in the sense of (3.60), (3.61). The number of equivalence classes in $\Upsilon_{j,q}$ w.r.t. $\sim$ can be bounded using a simple counting argument.
Proposition 3.34. The number of equivalence classes in $\mathcal{T}_{j,q}$ satisfies the bound

$$|\mathcal{T}_{j,q}| \sim \leq (M+1)^q(j+1)^q.$$  

Proof. Since every equivalence class contains exactly one element with non-decreasing entry sequence, we need to count (via (3.65)) the distinct solutions $n = (n_1, n_2, \ldots, n_M) \in \mathbb{N}^M$ of $n_1 + n_2 + \ldots + n_M = j$ with $n_m \geq 1$ for less than $q$ values of $1 \leq m \leq M$. We obtain

$$|\mathcal{T}_{j,q}| \sim = \sum_{i=0}^{q-1} \binom{M-i}{j-1} \leq (M+1)^q(j+1)^q.$$  

\[ \square \]

3.4.3 Complexity

The computational effort required by Algorithm 4 with sparsification index sets given by (3.62),

$$\mathcal{T}_j := \mathcal{T}_{j,q} \ \forall j \in \mathbb{N},$$  

and with $q \in \mathbb{N}$ to be chosen later follows directly from Propositions 3.32 and 3.34.

Proposition 3.35. Algorithm 4 consists in solving at most $(j+j'+1)^q+1(M+1)^q$ deterministic diffusion problems at FE discretization level $L$.

Combining Propositions 3.27, 3.33, 3.35 we obtain the following error and complexity estimate for Algorithm 5 below with exact input data,

$$h_m := M^{1+j+j'+1}(X_m, f) \quad \forall m \in \mathcal{T}_{j'+j}, \forall 0 \leq j \leq n.$$  

and sparsification index sets given by (3.72). We denote in the following the output of Algorithm 5 below with exact input data (3.73) by

$$M_{T,n}(r_M, u_M, L).$$  

Proposition 3.36. If $f \in L^1(\Omega, H^{-1}(D))$ and Assumptions 3.1, 3.18 hold, then for the solution $u_{M,L} \in L^1(\Omega, V_L)$ to (3.17),

$$\Delta_{e_M(\omega)} u_{M,L}(\omega) = f(\omega) \quad \text{in } V^*_L \quad \text{P.a.e. } \omega \in \Omega$$  

it holds (w.l.o.g. $c_a \geq 1$)

$$\|M'(r_M, u_{M,L}) - M_{T,n}(r_M, u_{M,L})\|_{A(U', H^{-1}(D))} \leq c_a\|f\|_{L^1(\Omega, H^{-1}(D))} \|r_M\|^n_{L^\infty(U, \Omega)} \exp(-c_1n) + c_f^{j+n}\exp(-c_2q^{j+1}).$$  

Moreover, $M_{T,n}(r_M, u_{M,L})$ can be computed starting from $M_{T}(r_M, f)$ for all $0 \leq j \leq n$ via Algorithm 5 below at the cost of solving at most $(j'+n+1)^q+2(M+1)^q$ deterministic diffusion problems in $D$ at FE level $L$. 

3.5 Mean Field Computation

From Propositions 3.17 and 3.36 with \( j' = 0 \) we obtain

**Proposition 3.37.** If \( f \in L^1(\Omega, H^{-1}(D)) \) and Assumptions 3.1, 3.18 hold, then

\[
\mathcal{M}^j(u_{M,L}) - \mathcal{M}^j_n(u_{M,L}) \|_{H_j^0(D)} \leq c_\alpha \| f \|_{L^1(\Omega, H^{-1}(D))} (\exp(-c_1, a, n) + c_\alpha^a \exp(-c_2, q^{1+\alpha})).
\]

(3.77)

If, in addition, \( f \in L^1(\Omega, H^{-1+s}(D)) \) and Assumption 3.12 holds too, then

\[
\mathcal{M}^j(u) - \mathcal{M}^j_n(u_{M,L}) \|_{H_j^0(D)} \leq c_\alpha, a \| f \|_{L^1(\Omega, H^{-1+s}(D))} (\exp(-c_1, M, a) + \Phi(N, s) + \exp(-c_1, a, n) + c_\alpha^a \exp(-c_2, q^{1+\alpha})).
\]

(3.78)

for any \( M \geq M_r, n, L, q \in \mathbb{N} \).

Moreover, \( \mathcal{M}^j_n(u_{M,L}) \) can be computed starting from \( \mathcal{M}^j_n(\tau_M, f) \) for all \( 0 \leq j \leq n \) via Algorithm 3 at the cost of solving at most \((n + 1)^{q+2}(M + 1)^q\) deterministic diffusion problems in \( D \) at FE level \( L \).

Note that the four terms on the r.h.s. of (3.78) correspond to the four discretization steps (modelling assumption, FEM in \( D \), truncation of Neumann series, and mixed moment sparsification for the pair \((\tau_M, f)\) respectively). Balancing the discretization errors in (3.78) by the (unique up to multiplicative constants) appropriate choice of parameters

\[
M \sim L^{1/\alpha}, \quad n \sim L, \quad q \sim L^{1/(1+\alpha)}
\]

(3.79)

and comparing with the computational effort estimate obtained in Proposition 3.37 we deduce (assuming also exact integration for all r.h.s.'s in step 3: of Algorithm 4),

**Theorem 3.38.** Under Assumptions 3.1 (well-posedness), 3.12 (uniform data regularity), 3.18 (analytic perturbation), the mean field of a random solution to (1.10) with \( f \in L^1(\Omega, H^{-1+s}(D)) \) is numerically computable using Algorithm 5 at nearly the same cost as one deterministic diffusion problem in \( D \).

If the available deterministic solver in \( D \) needs \( O(\varepsilon^{-\beta}) \) floating point operations and memory to reach an accuracy \( \varepsilon \) for one deterministic diffusion problem with coefficient \( c \) and \( H^s \) regular source term (usually \( \beta = 1/\delta + o(1) \) with \( \delta = \min(s, p)/d \)), then Algorithm 5 requires at most \( O(\varepsilon^{-\beta - o(1)}) \) operations and memory to achieve the same accuracy \( \varepsilon \to 0 \) for \( \mathcal{M}^j(u) \) in the case of stochastic data. Algorithm 5 has therefore nearly optimal complexity in the sense of Definition 1.1.

**Proof.** Choosing the discretization parameters \( M, n, q \) as in (3.79), the number of deterministic diffusion problems in \( D \) to be solved at FE level \( L \) (and given in Proposition 3.37) admits then (up to a multiplicative constant) the upper bound

\[
L^{O(L^{1/(1+\alpha)})} \leq \exp(O((\log N)_L^{1/(1+\alpha)} \log \log N_L)) \lesssim N_L^{o(1)} \quad \text{as } L \to \infty,
\]

(3.80)
which concludes the proof. □

Algorithms 4, 5 turn out to be also amenable to (almost full) parallelization.

**Remark 3.39.** The computation in step 1: of Algorithm 4 is fully parallelizable w.r.t. \( m \in T_{j'+j} \). The computation in step 3: of Algorithm 4 is also fully parallelizable w.r.t. \( m \in 
\Pi_{j'+j,k-1}(T_{j'+j}) \), but sequential w.r.t. \( k = j' + j : -1 : j' + 1 \) (that is, \( j \) serial steps are to be taken). Note that, due to (3.79), \( j \leq n \sim L \sim -\log \varepsilon \), where \( \varepsilon \) is the prescribed accuracy to be reached in the mean field/higher order moment computation.

Further, the computation in step 1: of Algorithm 5 is also fully parallelizable w.r.t. \( j = 0 : 1 : n \).

**Remark 3.40.** All results and algorithms derived for the mean field computation can be easily generalized to the case of arbitrary first order statistics (weighted expectations, obtained by replacing the probability measure \( P \) by a new one \( Q \) such that \( dQ/dP \) exists in \( L^\infty(\Omega) \) and it is known).

### 3.6 Computation of Higher Order Moments

Based on the results obtained in Chapter 2 (stochastic source term), on the mean field computation Algorithm 5 and under the regularity Assumption 3.18 we show that also higher order moments of the solution \( u \) to (1.10) can be computed in almost the same complexity as the solution of one deterministic diffusion problem in \( D \).

We formulate first a continuous moment problem which we subsequently discretize and decompose as direct sum of several mean field problems. These are in turn solved using Algorithm 5.

**Remark 3.41.** In the following we discuss (for simplicity and since correlation and variance kernels are often more relevant than higher order moments) only the computation of the second order moment \( (k = 2) \). Corresponding results hold for all higher order moments and can be derived in a similar manner.

For \( f \in L^0(\Omega, H^{-1}(D)) \) we recall that the diffusion problem (1.10) with stochastic data consists in finding \( u \in L^0(\Omega, H^1_0(D)) \) such that

\[
\Delta_{u(\omega)} u(\omega) = f(\omega) \quad \text{in } H^{-1}(D) \quad P\text{-a.e. } \omega \in \Omega. \tag{3.81}
\]

After truncating the fluctuation expansion (3.6) at level \( M \) (see also (3.11)) we are left with the problem (3.12) of finding \( u_M \in L^0(\Omega, H^1_0(D)) \) such that

\[
\Delta_{u_M(\omega)} u_M(\omega) = f(\omega) \quad \text{in } H^{-1}(D) \quad P\text{-a.e. } \omega \in \Omega. \tag{3.82}
\]

From Proposition 3.9 we immediately obtain
Proposition 3.42. If Assumption 3.4 holds, then for any $M > M_r$
\[ \| u(\omega) \otimes u(\omega) - u_M(\omega) \otimes u_M(\omega) \|_{H^1(D^2)} \leq c_{e,r,a} \exp(-c_1 r M^a) \| f(\omega) \|_{H^{-1}(D)}^2 \text{ P-a.e. } \omega \in \Omega, \]
so that
\[ \| \mathcal{M}^2(u) - \mathcal{M}^2(u_M) \|_{H^1(D^2)} \leq c_{e,r,a} \exp(-c_1 r M^a) \| f \|_{L^2(\Omega,H^{-1}(D))}^2. \] (3.83)

We consider also the diffusion problem with constant coefficient 1 and stochastic right hand side $f \in L^0(\Omega, H^{-1}(D))$, whose solution we denote by $\bar{f} \in L^0(\Omega, H_0^1(D))$,
\[ \Delta \bar{f}(\omega) = f(\omega) \text{ in } H^{-1}(D) \text{ P-a.e. } \omega \in \Omega. \] (3.85)

We formulate next a continuous variational problem, to which $u_M(\cdot) \otimes u_M(\cdot) \in L^0(\Omega, H_0^1(D^2))$ is the unique solution. This problem will be in turn fully discretized. By $\text{Id}_{H_0^1(D)}$ we denote in the following the identity operator acting in $H_0^1(D)$.

Proposition 3.43. P-a.e. in $\Omega$ it holds,
\[ (\Delta \otimes \Delta)(\bar{f}(\omega) \otimes \bar{f}(\omega)) = f(\omega) \otimes f(\omega) \text{ in } H^{-1}(D) \otimes H^{-1}(D), \] (3.86)
\[ (\Delta_a M(\omega) \otimes \text{Id}_{H_0^1(D)})(u_M(\omega) \otimes \bar{f}(\omega)) = (\Delta \otimes \text{Id}_{H_0^1(D)})(\bar{f}(\omega) \otimes \bar{f}(\omega)) \text{ in } H^{-1}(D) \otimes H_0^1(D), \] (3.87)
\[ (\text{Id}_{H_0^1(D)} \otimes \Delta_a M(\omega))(u_M(\omega) \otimes u_M(\omega)) = (\text{Id}_{H_0^1(D)} \otimes \Delta)(u_M(\omega) \otimes \bar{f}(\omega)) \text{ in } H_0^1(D) \otimes H^{-1}(D). \] (3.88)

Proof. All three equations follow directly from (3.82) and (3.85) by tensorization.

3.6.1 Discretization

We introduce the sparse tensor product spaces in $D \times D$ by
\[ \tilde{V}_L := \text{Span } \{ V_l \otimes V_{l'} \mid 0 \leq l + l' \leq L \} \subset H_0^1(D^2). \] (3.89)

Under Assumption 2.11, the approximation property of the scale $(\tilde{V}_L)_{L \in \mathbb{N}}$ reads (see e.g. [Zen91], [ST03b])
\[ \min_{v \in \tilde{V}_L} \| u - v \|_{H_0^1(D^2)} \leq \hat{\Phi}(N_L, s) \| u \|_{H^{-1+s}(D^2)} \forall u \in H^{-1+s}(D^2) \cap H_0^1(D^2), \] (3.90)
where the functional $\hat{\Phi}$ can be explicitly written in terms of $\Phi$. For instance we have $\hat{\Phi}(N_L) := c_T (\log N_L)^{1/2} N_L^{-\delta}$ with $\delta := \min\{p, s\}/d$ for the FE spaces discussed in Example 2.12.

We further assume the existence of a wavelet family in the physical domain $D$, consistent with the FE space scale $(V_L)_{L \in \mathbb{N}}$. 

3.6. COMPUTATION OF HIGHER ORDER MOMENTS

Assumption 3.44. There exist a family $\Psi = \{\psi_{k,l}\}_{k\in \mathbb{N}} \subset H^1_0(D)$ and two constants $c_1, c_2 > 0$ such that any $u \in H^1_0(D)$ can be expanded as a convergent series,

$$u = \sum_{l \in \mathbb{N}} c_{k,l} \psi_{k,l} \in H^1_0(D) \quad (3.91)$$

and the following three conditions are fulfilled:

1. stability,

$$c_1 \sum_{l \in \mathbb{N}} |c_{k,l}|^2 \leq \| \sum_{l \in \mathbb{N}} c_{k,l} \psi_{k,l} \|_{H^1_0(D)}^2 \leq c_2 \sum_{l \in \mathbb{N}} |c_{k,l}|^2. \quad (3.92)$$

2. local support,

$|\Lambda_l| \sim 2^{d} l$ and for any $l, l'$, the set $\text{supp}(\psi_{k,l}) \cap \text{supp}(\psi_{k',l'})$ has nonempty interior for at most $\sim (l + l' + 1)^{d-1} \cdot 2^{\max(l,l')}$ values of $(k, k') \in \Lambda_l \times \Lambda_{l'}$.

3. consistency,

$$V_L = \text{span}\{\psi_{k,l} | 0 \leq l \leq L, k \in V_l\}.$$ 

Constructions of families $\Psi$ satisfying Assumption 3.44 are known for intervals ($D = [0, 1]$), hypercubes ($D = [0, 1]^d$) or polygonal domains (see e.g. [DS99]). Note that the stability condition ensures the existence of a scalar product $(\cdot, \cdot)_\Psi$ in $H^1_0(D)$ which is equivalent to the standard one and has the property that the wavelet family $\Psi := \{\psi_{k,l}\}_{k \in \mathbb{N}}$ is an ONB in $(H^1_0(D), (\cdot, \cdot)_\Psi)$.

We consider the discrete counterparts of (3.86), (3.87), (3.88) associated to the sparse FE space hierarchy $(\hat{V}_L)_{L \in \mathbb{N}}$, and consisting in finding $(\hat{f}_L), (u_L)_{M,L}, (uu)_{M,L} \in L^0(\Omega, \hat{V}_L)$ such that

$$(\Delta \otimes \Delta)((\hat{f}_L)_L(\omega)) = f(\omega) \otimes f(\omega) \quad \text{in } \hat{V}_L^*, \quad (3.93)$$

$$(\Delta_{a_M(\omega)} \otimes \text{Id}_{H^1_0(D)})((uu)_{M,L}(\omega)) = (\Delta \otimes \text{Id}_{H^1_0(D)})((\hat{f}_L)_L(\omega)) \quad \text{in } \hat{V}_L^*, \quad (3.94)$$

$$(\text{Id}_{H^1_0(D)} \otimes \Delta_{a_M(\omega)}((uu)_{M,L}(\omega)) = (\text{Id}_{H^1_0(D)} \otimes \Delta)((\hat{f}_L)_L(\omega)) \quad \text{in } \hat{V}_L^*. \quad (3.95)$$

More precisely, by (3.94) we understand the variational formulation which in the second variable is associated to the scalar product (bilinear form) $(\cdot, \cdot)_\Psi$, and similarly for (3.95).

Proposition 3.45. Under Assumption 3.12, it holds, $P$-a.e. in $\Omega$ and for all $M \geq M_\tau$ (here $\text{Id} := \text{Id}_{H^1_0(D^2)}$),

$$\|u_M(\omega) \otimes u_M(\omega) - (uu)_{M,L}(\omega)\|_{H^1_0(D^2)} \leq c_{a,\Psi}(\|\text{Id} - P_{\hat{V}_L})(\hat{f}(\omega) \otimes \hat{f}(\omega))\|_{H^1_0(D^2)} + \|\text{Id} - P_{\hat{V}_L})(u_M(\omega) \otimes u_M(\omega))\|_{H^1_0(D^2)} \quad (3.96)$$

so that if also Assumption 3.12 is satisfied, then $P$-a.e. $\omega \in \Omega$

$$\|u_M(\omega) \otimes u_M(\omega) - (uu)_{M,L}(\omega)\|_{H^1_0(D^2)} \leq c_{a,\Psi} \hat{\Phi}(N_L) \|f(\omega)\|_{H^{-1+\epsilon}(D)}^2 \quad (3.97)$$
and
\[ \|M^2(u_M) - M^1((uu)_{M,L})\|_{H^1(D^2)} \leq c_a,\Phi(\tilde{N}_L)\|f\|^2_{L^2(\Omega,H_{-\delta}(D))}. \] (3.98)

Recall that \( \hat{\Phi}(\tilde{N}_L) := c_T(\log N_L)^{1/2}N_L^{-\delta} \) with \( \delta := \min\{p,s\}/d \) for the FE spaces discussed in Example 2.12.

**Proof.** The P-a.e. uniform ellipticity of the bilinear forms giving (the l.h.s.'s of) (3.86), (3.87), (3.88) for \( M \geq M_r \) ensures the quasi-optimality of the corresponding FE solutions in (3.93), (3.94), (3.95). We thus have, P-a.e. in \( \Omega, \)
\[ \|u_M(\omega) \otimes u_M(\omega) - (uu)_{M,L}(\omega)\|_{H_0^1(D^2)} \leq c_a,\Phi(\|Id - P_{V_L}(u_M(\omega) \otimes u_M(\omega))\|_{H_0^1(D^2)} + \\
+ \|u_M(\omega) \otimes f(\omega) - (u_f)_{M,L}(\omega)\|_{H_0^1(D^2)} \) (3.99)
\[ \|u_M(\omega) \otimes f(\omega) - (u_f)_{M,L}(\omega)\|_{H_0^1(D^2)} \leq c_a,\Phi(\|Id - P_{V_L}(u_M(\omega) \otimes f(\omega))\|_{H_0^1(D^2)} + \\
+ \|f(\omega) \otimes f(\omega) - (f_f)_{L}(\omega)\|_{H_0^1(D^2)} \) (3.100)
\[ \|f(\omega) \otimes f(\omega) - (f_f)_{L}(\omega)\|_{H_0^1(D^2)} \leq \|Id - P_{V_L}(f(\omega) \otimes f(\omega))\|_{H_0^1(D^2)}. \) (3.101)

Inserting (3.101) in (3.100) and then the resulting estimate (3.100) in (3.99) we obtain (3.96). Further, (3.97) follows from (3.96) and the approximation property (3.90) of the sparse tensor product spaces. Integrating (3.97) over \( \Omega \) we obtain (3.98). \( \square \)

### 3.6.2 Algorithms

We start by decomposing equations (3.94), (3.95) into simpler, mean field problems using the wavelet family \( \Psi := (\psi_{k,l})_{k \in \Lambda_L} \). Since \( (f_f)_{L}, (u_f)_{M,L}, (uu)_{M,L} \in L^0(\Omega, V_L) \), we can write
\[ (f_f)_L = \sum_{k \leq L, k \in \Lambda_L} \alpha_{k,l} \otimes \psi_{k,L-l} \] (3.102)
\[ (u_f)_{M,L} = \sum_{k \leq L, k \in \Lambda_L} \beta_{k,l} \otimes \psi_{k,L-l} \] (3.103)
\[ = \sum_{k \leq L, k \in \Lambda_L} \psi_{k,L-l} \otimes \gamma_{k,l} \] (3.104)
\[ (uu)_{M,L} = \sum_{k \leq L, k \in \Lambda_L} \psi_{k,L-l} \otimes \delta_{k,l} \] (3.105)

where \( \alpha_{k,l}, \beta_{k,l}, \gamma_{k,l}, \delta_{k,l} \in L^0(\Omega, V_L) \) for all \( 0 \leq l \leq L, k \in \Lambda_L \).

For future use we note here that the family \( (\gamma_{k,l})_{k \leq L, k \in \Lambda_L} \) can be obtained from the family \( (\beta_{k,l})_{k \leq L, k \in \Lambda_L} \) by a rearrangement of its coefficients in the \( V_L \) basis \( (\psi_{k,l})_{k \leq L, k \in \Lambda_L} \). The proof follows by simple manipulations of (3.103) and (3.104).
3.6. COMPUTATION OF HIGHER ORDER MOMENTS

Proposition 3.46. It holds
\[ \gamma_{k,l}(\omega) = \sum_{\omega \in \Omega} \langle \beta_{k',l',\nu}(\omega), \psi_{k,l-\nu} \rangle_{\psi_{k',l'}} \quad P\text{-a.e. } \omega \in \Omega. \tag{3.106} \]

Using the decompositions (3.102), (3.103), (3.104), (3.105) in the variational formulation of (3.94), (3.95) we obtain

Proposition 3.47. Equation (3.94) is equivalent to
\[ \Delta_{\alpha_M}(\omega) \beta_{k,l}(\omega) = \Delta_{\alpha_{k,l}}(\omega) \quad \text{in } V^*_1 \quad \forall 0 \leq l \leq L, k \in \Lambda_{L-l} \quad P\text{-a.e. } \omega \in \Omega. \tag{3.107} \]

Similarly, equation (3.95) is equivalent to
\[ \Delta_{\alpha_M}(\omega) \delta_{k,l}(\omega) = \Delta_{\gamma_{k,l}}(\omega) \quad \text{in } V^*_1 \quad \forall 0 \leq l \leq L, k \in \Lambda_{L-l} \quad P\text{-a.e. } \omega \in \Omega. \tag{3.108} \]

Keeping in mind the results of section 3.4, let us explain how the correlation computation algorithm based on the discretization step given by (3.93), (3.94), (3.95) should work. The correlation kernel \( \mathcal{M}^1((u_w)_{M,L}) \) can be obtained using (3.105) from the mean fields \( \mathcal{M}^1(\delta_{k,l}) \), which in turn follow in view of (3.108) from the mixed moments of \( (r_M, \beta_{k,l}) \). By rearrangement via (3.106), the mixed moments of \( (r_M, \gamma_{k,l}) \) are expressed in terms of the mixed moments of \( (r_M, \beta_{k,l}) \), which follow then from the mixed moments of \( (r_M, \alpha_{k,l}) \). Finally, the mixed moments of \( (r_M, \alpha_{k,l}) \) can be obtained, due to (3.102), from the mixed moments of \( (r_M, f \otimes f) \) (or of \( (r_M, f \otimes f) \)). The algorithm will therefore have to perform all these steps in reverse order, starting from the available information, the mixed moments of the pair \( (r_M, f \otimes f) \).

In the following we use the notation
\[ (\Delta \otimes \Delta)^{-1} : H^{-1}(D) \otimes H^{-1}(D) \rightarrow \tilde{V}_L \tag{3.109} \]

for the discrete solution operator associated to the bilinear form of \( \Delta \otimes \Delta \) (coercive on \( H^1_0(D^2) \)) and the sparse FE space \( \tilde{V}_L \). Further, we denote by
\[ \iota_{\psi,\phi} : H^1_0(D) \rightarrow \mathbb{C}, \quad \iota_{\psi,\phi}(\xi) := \langle \xi, \psi \rangle_{\psi} \quad \forall \xi \in H^1_0(D) \tag{3.110} \]

the contraction in the Hilbert space \( (H^1_0(D), \langle \cdot, \cdot \rangle_{\psi}) \) with vector \( \psi \in H^1_0(D) \).

With these notations, simple algebra shows the connection between the mixed moments of the pair \( (r_M, f \otimes f) \) and those of \( (r_M, \alpha_{k,l}) \).

Proposition 3.48. In \( \mathcal{A}(U^1, H^1_0(D)) \) it holds
\[ \mathcal{M}^{(2,1)}_{T} (r_M, \alpha_{k,l}) = \sum_{m \in T} g_{m,\alpha_{k,l}} \otimes \phi_m \tag{3.111} \]

where
\[ g_{m,\alpha_{k,l}} = (\text{Id}_{H^1_0(D)} \otimes \iota_{\psi,\phi_{k,l}})(\Delta \otimes \Delta)^{-1} \mathcal{M}^{1,1} (X_m, f \otimes f) \in V_1 \tag{3.112} \]

satisfy the symmetry condition w.r.t. \( m \) (invariance under entry permutations).
Proof. Using (3.102) and the linearity of the mixed moments of order \((j, 1)\) in the last argument we write

\[
M_{1}^{(j,1)}(rM, (f\bar{f})L) = \sum_{\varnothing \subseteq L, k \in A_{L-1}} M_{1}^{(j,1)}(rM, \alpha_{k,i}) \otimes \psi_{k,L-1}.
\] (3.113)

But (3.93) shows that

\[
M_{1}^{(j,1)}(rM, (f\bar{f})L) = \sum_{m \in T_{j}} \left( \Delta \otimes \Delta \right)^{-1}_{L} M_{1}^{(j+1)}(X_{m}, f \otimes \bar{f}) \otimes \phi_{m},
\] (3.114)

so that the conclusion follows by applying \(\text{Id}_{H_{0}^{1}(D)} \otimes \psi_{k,L-1} \otimes \text{Id}_{L^{\infty}(U)}\) to (3.114) and (3.113). \(\square\)

Noting that, due to (3.93) and (3.102), \(f \in L^{2}(\Omega, H_{0}^{1}(D))\) ensures \(\alpha_{k,i} \in L^{1}(\Omega, H_{0}^{1}(D))\) boundedly, we apply Proposition 3.36 to the equation (3.107).

**Proposition 3.49.** Under Assumptions 3.1, 3.18, \(M_{1}^{(j,1)}(rM, \beta_{k,i})\) can be computed starting from \(M_{1}^{(j,1)}(rM, \gamma_{k,i})\) for all \(0 \leq j \leq n\) via Algorithm 5 at the cost of solving at most \((j' + n + 1)q^{j+2}(M+1)^{q}\) deterministic diffusion problems in \(D\) at FE level \(l\). The error estimate reads

\[
\|M_{1}^{(j,1)}(rM, \beta_{k,i}) - M_{1}^{(j,1)}(rM, \gamma_{k,i})\|_{A(U^{j'}H_{0}^{1}(D))} \leq \sum_{m \in T_{j}} c_{a} \|\alpha_{k,i}\|_{L^{1}(\Omega,H_{0}^{1}(D))}(\|rM\|_{L^{\infty}(U^{j'}D)}) \exp(-c_{1,a}n) + c_{a}^{j+n+1} \exp(-c_{2,q^{j+1}+a}).
\] (3.115)

The next step of the algorithm, passing from the mixed moments of the pair \((rM, \beta_{k,i})\) to the mixed moments of \((rM, \gamma_{k,i})\), can be analyzed using Proposition 3.46 and the linearity of the approximate mixed moment \(M_{1}^{(j,1)}\) in the last argument (see (3.59)).

**Proposition 3.50.** If

\[
M_{1}^{(j,1)}(rM, \beta_{k,i}) = \sum_{m \in T_{j'}} g_{m,\beta_{k,i}} \otimes \phi_{m},
\] (3.116)

with \(g_{m,\beta_{k,i}} \in V_{l}\), then

\[
M_{1}^{(j,1)}(rM, \gamma_{k,i}) = \sum_{m \in T_{j'}} g_{m,\gamma_{k,i}} \otimes \phi_{m}
\] (3.117)

where \(g_{m,\gamma_{k,i}} \in V_{l}\) is given by

\[
g_{m,\gamma_{k,i}} = \sum_{\varnothing \subseteq L, k \in A_{L-1}} (g_{m,\beta_{k',L-1}', \psi_{k,L-1}'} \otimes \psi_{k',L-1}').
\] (3.118)

Moreover, the following error estimate holds,

\[
\|M_{1}^{(j,1)}(rM, \gamma_{k,i}) - M_{1}^{(j,1)}(rM, \beta_{k,i})\|_{A(U^{j'}H_{0}^{1}(D))} \leq \sum_{m \in T_{j'}} c_{a} \|f\|_{L^{2}(\Omega,H^{-1}(D))}(\|rM\|_{L^{\infty}(U^{j'}D)}) \exp(-c_{1,a}n) + c_{a}^{j+n} \exp(-c_{2,q^{j+1}+a}).
\] (3.119)
Proof. From (3.106) we obtain

\[
\mathcal{M}_{T,n}^{(j',1)}(r_M, \gamma_{k,l}) = \sum_{0 \leq l' \leq l} \mathcal{M}_{T,n}^{(j',1)}(r_M, (\beta_{k',L-l'}(\cdot), \psi_{k,k-L-l'}(\cdot), \psi_{k',l'})
\]

\[
= \sum_{0 \leq l' \leq l} \left( \mathbf{I}_{L=\infty(D;l')} \otimes \mathcal{M}_{T,n}^{(j',1)}(r_M, \beta_{k',L-l'}(\cdot) \otimes \psi_{k',l'})
\]

\[
= \sum_{0 \leq l' \leq l} \sum_{k' \in \Lambda} \left( \mathbf{I}_{L=\infty(D;l')} \otimes \mathcal{M}_{T,n}^{(j',1)}(r_M, \beta_{k',L-l'}(\cdot) \otimes \psi_{k',l'})
\]

from which (3.118) follows by interchanging the summations. To check (3.119) we first note that \( \gamma_{k,l} \) solves, in view of (3.106),

\[
\Delta_{\alpha, \gamma_{k,l}}(\omega) \gamma_{k,l}(\omega) = \sum_{0 \leq l' \leq l} (\beta_{k',L-l'}(\omega), \psi_{k,k-L-l'}(\cdot), \psi_{k',l'}) \gamma_{k,l}(\omega) \gamma_{k,l}(\omega) \in V_1^* \text{ P-a.e. } \omega \in \Omega. \quad (3.120)
\]

Next we observe that the \( L^1(\Omega, H^{-1}(D)) \) norm of the r.h.s. in (3.120) can be estimated from above by

\[
c_a \int \left( \sum_{0 \leq l' \leq l} \| \beta_{k',L-l'}(\omega) \|_{H_0^1(D)} \right)^{1/2} dP(\omega) \leq c_a \| (u)_{M,L} \|_{L^1(\Omega, H_0^1(D^2))} \leq c_a \| f \|_{L^2(\Omega, H^{-1}(D))},
\]

so that (3.119) follows from Proposition 3.36 applied to the equation (3.120).

Finally we investigate the computation of \( \delta_{k,l} \) solution to (3.108). We apply the \( \Upsilon \) version of Proposition 3.27 (see also Remark 3.31) to equation (3.108) using as input data for Algorithm 5 the approximate moments

\[
\mathcal{M}_{app,\Upsilon}^{(j',1)}(r_M, \Delta \gamma_{k,l}) := \mathcal{M}_{T,n}^{(j',1)}(r_M, \Delta \gamma_{k,l})
\]

and obtain

**Proposition 3.51.** Under Assumptions 3.1, 3.18, an approximate mean field

\[
\mathcal{M}_{app,\Upsilon}^{1}(\delta_{k,l}) := \text{Algorithm 5 } [M, n', 0, (\mathcal{M}_{T,n}^{(j',1)}(r_M, \Delta \gamma_{k,l}))_{0 \leq l' \leq l'}]
\]

can be computed at the cost of solving at most \((n' + 1)^q + (M + 1)^q\) deterministic diffusion problems in \( D \) at FE level \( l \) and for an accuracy (w.l.o.g. \( c_a \| r_M \|_{L^\infty(U \times \Omega)} \geq 1 \))

\[
\| M_1(\delta_{k,l}) - \mathcal{M}_{app,\Upsilon,n,n'}(\delta_{k,l}) \|_{H_0^1(D)} \leq
\]

\[
\leq c_a \| f \|_{L^2(\Omega, H^{-1}(D))} (\exp(-c_1a n') + \| r_M \|_{L^\infty(U \times \Omega)} \exp(-c_1a n') + c_a^{n+n} \exp(-c_2a r^{1+n})),
\]

(3.123)

Due to the orthogonal decomposition (3.105) we then have, with \( \hat{N}_L := \text{dim} \hat{V}_L \),
Corollary 3.52. Under Assumptions 3.1, 3.18 it holds

\[ \|M^i((uu)_{M,L}) - \sum_{0 \leq i \leq L} \psi_{k,l-i} \otimes M^i_{\text{app},T,n,n'(\delta_{k,i})}\|_{H^1(\mathcal{D})} \leq \]

\[ \leq c_a \|f\|^2_{L^2(\Omega, H^{-1}(\mathcal{D}))} \hat{N}_L^{i/2}(\exp(-c_{1,a}n') + \|r_M\|_{L^\infty(U \times \Omega)} \exp(-c_{1,a}n + c_{n+n} \exp(-c_2,q^{1+\alpha})). \]

(3.124)

We summarize more formally in Algorithm 6 the second moment computation we have developed and analyzed in this section. The total error estimate follows by collecting all discretization errors, as given by Propositions 3.42, 3.45 and Corollary 3.52.

Proposition 3.53. It holds

\[ \|M^2(u) - \sum_{0 \leq i \leq L} \psi_{k,l-i} \otimes M^i_{\text{app},T,n,n'}(\delta_{k,i})\|_{H^1(\mathcal{D})} \leq \]

\[ \leq c_{a,r} \|f\|^2_{L^2(\Omega, H^{-1}(\mathcal{D}))} \exp(-c_{1,r}M^\alpha) + \hat{\Phi}(N_L) + \hat{N}_L^{i/2}(\exp(-c_{1,a}n') + \]

\[ + \|r_M\|_{L^\infty(U \times \Omega)} \exp(-c_{1,a}n + c_{n+n} \exp(-c_2,q^{1+\alpha})). \]

(3.125)

The complexity of Algorithm 6 can be estimated using Propositions 3.49 and 3.51.

Proposition 3.54. Step 1 of Algorithm 6 consists in solving \(|\Theta_{n' + n}/ \sim | + |\Theta_{n' + n-1}/ \sim | + \cdots + |\Theta_0/ \sim | \) second moment diffusion problems in \(D \times D\) with constant coefficient and in the sparse FE space \(\hat{V}_L\).

Steps 2 and 4 of Algorithm 6 consist in solving at most \(|\lambda_{l-1}(n' + n + 1)^{2+2}(M + 1)^q\) deterministic diffusion problems in \(D\) with coefficient \(e\) and in the FE space \(V_l\) for all \(0 \leq l \leq L\).

Step 3 of Algorithm 6 consists only of a sorting procedure applied to the coefficient sets of \((g_m, \delta_{k,l})_{0 \leq i \leq L}\) in the wavelet basis \(\Psi\) (compare (3.118)).

Proof. The complexity estimate of step 1 follows from the moment symmetry w.r.t. \(m\), whereas those of steps 2 and 4 are direct consequences of Proposition 3.36. \(\square\)

Note that a more explicit complexity estimate of step 1: in Algorithm 6 follows from the main result in Chapter 2, that we recall next.

Theorem 3.55. Under Assumption 2.15 and if \(f \in L^2(\Omega, H^{-1+\epsilon}(\mathcal{D}))\), each problem in step 1 of Algorithm 6 is numerically solvable for any \(L \geq 1\) at a cost of

\[ c_T,\phi((log N_L)^{2q+2} N_L \] (3.126)

floating point operations, with

\[ c_T((log N_L)N_L \] (3.127)

memory, for an accuracy of

\[ c_{a,a,T}(log N_L)^{1/2} N_L^{-\delta}, \] (3.128)

where \(\delta = \min\{p, s\}/d\).
Balancing the discretization errors in (3.125) by the choice of parameters

\[ M \sim L^{1/\alpha}, \quad n' \sim L, \quad n \sim L, \quad q \sim L^{1/(1+\alpha)} \]  

(3.129)

and comparing with the computational effort estimate obtained in Proposition 3.54 we deduce (assuming also exact integration for the r.h.s. of all deterministic problems to be solved in Algorithm 6),

**Theorem 3.56.** Under Assumptions 3.1 (well-posedness), 3.12 (uniform data regularity), 3.18 (analytic perturbation) and 2.15 (existence of $H^3_0(D)$-stable wavelet family), the second order moment (correlation) of a random solution to (1.10) with $f \in L^2(\Omega, H^{-1+s}(D))$ is computable using Algorithm 6 at nearly the same cost as one deterministic diffusion problem in $D$.

More precisely, if the available wavelet based solver in $D$ needs $O(\varepsilon^{-\beta})$ floating point operations and memory to reach an accuracy $\varepsilon$ for one deterministic diffusion problem with coefficient $\varepsilon$ and $H^s$ regular source term (usually $\beta = 1/\delta + o(1)$ with $\delta = \min(s, p)/d$), then Algorithm 6 requires at most $O(\varepsilon^{-\beta-o(1)})$ operations and memory to achieve the same accuracy $\varepsilon \to 0$ for $M^2(u)$ in the case of stochastic data. Algorithm 6 has therefore nearly optimal complexity in the sense of Definition 1.1.

**Proof.** Due to Propositions 3.34, 3.54 and Theorem 2.28, the computational effort of step 1 consists in at most $NL^L$ operations as $L \to \infty$, by the same argument used in (3.80). This argument shows also that the cost of steps 2 and 4 admits (up to a multiplicative constant) the upper bound

\[ \sum_{l=0}^{L} |A_{L-l}| N_{L}^{\sigma(1)} N_{l}^{1+\sigma(1)} \lesssim N_{L}^{1+\sigma(1)} \quad \text{as} \quad L \to \infty, \]  

(3.130)

for an accuracy of $\Phi(N_L) \sim N_{L}^{-\delta+o(1)}$ (compare Theorem 3.55).

**Remark 3.57.** All solution operations contained in steps 1; 2; 3; 4: of Algorithm 6 are fully parallelizable w.r.t. the indices $l, j, j'$ and $k \in A_{L-l}$. \qed
input: • $M, j, j', L$
  • approximation of $\mathcal{M}^{(j',1)}(r_M, f)$ of the form

$$\mathcal{M}^{(j',1)}_{\text{app}, T}(r_M, f) = \sum_{m \in \mathcal{T}_{j'+j}} h_m \phi_m \quad \text{with } h_m \in H^{-1}(D)$$

% e.g. $h_m := \mathcal{M}^{1, j'+j+1}(X_m, f)$ for exact sparsified representation of $\mathcal{M}^{(j',1)}_{T}(r_M, f)$ (see (3.73))
  • solver of diffusion problem in $D$ with deterministic coefficient $e$ and FE space $V_L$

output: • approximation of $\mathcal{M}^{(j',1)}(r_M, v_{j,L})$ of the form

$$\mathcal{M}^{(j',1)}_{\text{app}, T}(r_M, v_{j,L}) = \sum_{m \in \mathcal{T}_{j'+j}} g_m \phi_m$$

with $g_m \in V_L$

1 compute for all $m \in \mathcal{T}_{j'+j}$

$$g_m := \Delta_{e,L}^{-1} h_m \in V_L$$

2 for $k = j' + j : -1 : j' + 1$ do

3  for all $m \in \mathcal{P}_{j'+j,k-1}(\mathcal{T}_{j'+j})$

$$g_m := \Delta_{e,L}^{-1} \left( \text{div} \sum_{m' \in \mathcal{P}_{j'+j,k}(\mathcal{T}_{j'+j})} \sum_{\mathcal{P}_{j'+j,k-1}(m') = m} \phi_{m'} \nabla g_{m'} \right) \in V_L$$

4 end

5 return

$$\mathcal{M}^{(j',1)}_{\text{app}, T}(r_M, v_{j,L}) := \sum_{m \in \mathcal{T}_{j'+j}} g_m \phi_m$$

Algorithm 4: $\mathcal{M}^{(j',1)}_{\text{app}, T}(r_M, v_{j,L}) = \text{Algorithm 4} \left[ M, j, j', L, \mathcal{M}^{(j'+j,1)}_{\text{app}, T}(r_M, f) \right]$
input: • $M$, $n$, $j'$, $L$

- approximations of $\mathcal{M}^{(j'+j,1)}(r_M, f)$ for all $0 \leq j \leq n$, of the form
  \[ \mathcal{M}^{(j'+j',1)}_{\text{app},T}(r_M, f) = \sum_{m \in \Gamma_{j'+j}} h_m \phi_m \quad \text{with} \quad h_m \in H^{-1}(D) \]

  e.g. $h_m := \mathcal{M}^{1,1}_{j'+j+1}(X_m, f)$ for exact sparsified representation of $\mathcal{M}^{(j'+j,1)}_{T}(r_M, f)$ (see (3.73))

- solver of diffusion problem in $D$ with deterministic coefficient $e$ and FE space $V_L$

output: • approximation of $\mathcal{M}^{(j',1)}(r_M, u_{M,L})$ of the form
  \[ \mathcal{M}^{(j',1)}_{\text{app},T,n}(r_M, u_{M,L}) = \sum_{m \in \Pi_{j'+j}} g_m \phi_m \]
  with $g_m \in V_L$

1 compute for $j = 0 : 1 : n$

  \[ \mathcal{M}^{(j',1)}_{\text{app},T}(r_M, u_{j,L}) := \text{Algorithm 4} \left[ M, j, j', L, \mathcal{M}^{(j'+j,1)}_{\text{app},T}(r_M, f) \right] \]

2 compute and return

  \[ \mathcal{M}^{(j',1)}_{\text{app},T,n}(r_M, u_{M,L}) := \sum_{j=0}^{n} \mathcal{M}^{(j',1)}_{\text{app},T}(r_M, u_{j,L}) \]

Algorithm 5: $\mathcal{M}^{(j',1)}_{\text{app},T,n}(r_M, u_{M,L}) = \text{Algorithm 5} \left[ M, n, j', L, (\mathcal{M}^{(j'+j,1)}_{\text{app},T}(r_M, f))_{0 \leq j \leq n} \right]$
input: • \( M, n, n', L \)
  • approximations of \( \mathcal{M}^{(j'+j,1)}(r_M, f \otimes f) \) for all \( 0 \leq j \leq n, 0 \leq j' \leq n' \) of the form

\[
\mathcal{M}^{(j'+j,1)}(r_M, f \otimes f) = \sum_{m \in T_{j'+j}} h_m \phi_m
\]

with \( h_m := \mathcal{M}^{1,j'+j+1}(X_m, f \otimes f) \in H^{-1}(D^2) \)

• solver of deterministic diffusion problem in \( D \) with coefficient \( e \) on FE space scale \( (V_l)_{0 \leq l \leq L} \)

• \( \Psi \)-wavelet based deterministic solver of second moment diffusion problem in \( D \times D \) with sparse FE space \( \tilde{V}_L \)

output: • approximation \( \mathcal{M}_\text{app,T,n,n'}^1((uu)_M,L) \in \tilde{V}_L \) of \( \mathcal{M}^2(u) \)

1 compute for all \( m \in T_{j'+j}, 0 \leq j \leq n, 0 \leq j' \leq n' \) and all \( 0 \leq l \leq L, k \in \Lambda_{L-l} \)

\[
g_{m,\alpha_k,l} := (1d_{\mathbb{H}_2(D)} \otimes \iota_{\Psi, \nu_{k,L-l}})(\Delta \otimes \Delta)^{-1}_L h_m \in V_l
\]

2 compute for all \( 0 \leq j' \leq n' \) and all \( 0 \leq l \leq L, k \in \Lambda_{L-l} \),

\[
\mathcal{M}^{(j',1)}_{T,n}(r_M, \beta_{k,l}) = \sum_{m \in T_{j'}} g_{m,\beta_{k,l}} \otimes \phi_m
\]

by

\[
\mathcal{M}^{(j',1)}_{T,n}(r_M, \beta_{k,l}) = \text{Algorithm 5} \left[ M, n, j', l, (\mathcal{M}^{(j'+j,1)}_{T}(r_M, \Delta \alpha_{k,l}))_{0 \leq j, \leq n} \right]
\]

3 compute for all \( 0 \leq j' \leq n' \) and all \( 0 \leq l \leq L, k \in \Lambda_{L-l} \) via (3.118),

\[
\mathcal{M}^{(j',1)}_{T,n}(r_M, \gamma_{k,l}) = \sum_{m \in T_{j'}} g_{m,\gamma_{k,l}} \otimes \phi_m
\]

4 compute for all \( 0 \leq l \leq L, k \in \Lambda_{L-l} \),

\[
\mathcal{M}^{(0,1)}_{\text{app,T,n,n'}}(r_M, \delta_{k,l}) = \text{Algorithm 5} \left[ M, n', 0, l, (\mathcal{M}^{(j',1)}_{T,n}(r_M, \Delta \gamma_{k,l}))_{0 \leq j' \leq n'} \right]
\]

5 compute and return

\[
\mathcal{M}^2_{\text{app,T,n,n'}}((uu)_M,L) := \sum_{0 \leq l \leq L, k \in \Lambda_{L-l}} \psi_{k,L-l} \otimes \mathcal{M}^1_{\text{app,T,n,n'}}(\delta_{k,l})
\]

Algorithm 6: \( (\mathcal{M}^{(0,1)}_{\text{app,T,n,n'}}(\delta_{k,l}))_{0 \leq l \leq L, k \in \Lambda_{L-l}} = \text{Algorithm 6}[M, n, n', L, (\mathcal{M}^{(j'+j,1)}_{T}(r_M, f \otimes f))_{0 \leq j \leq n, 0 \leq j' \leq n'}] \)
Chapter 4

Conclusions and Open Issues

Classical perturbation algorithms for the moment computation of the solution to an elliptic problem with stochastic data have been further developed and investigated in terms of complexity. Statistical information about the data, needed as algorithm input, has been explicitly determined. It has been shown that, under essentially only one additional assumption, the spatial regularity of the fluctuation, the algorithms are, asymptotically in work vs. accuracy, more efficient than the standard MC/SG/WPC methods, and a priori nearly optimal (in the sense of Definition 1.1) complexity bounds have been proven. As in the case of a standard MC simulation, the computation is almost fully parallelizable and the implementation requires essentially only one routine to call any available, preexistent FE code for the corresponding deterministic elliptic problem and run it successively with different source terms.

We conclude by enumerating several possible extensions of the results presented in this thesis, along with some further open problems for the solution of which the techniques and ideas used here might be, in our opinion, relevant.

1. Relaxation of the fluctuation regularity condition \( r \in \mathcal{A}(\overline{D}, L^\infty(\Omega)) \) to only finite Sobolev regularity. We further conjecture that, in order that perturbation algorithms retain nearly optimal complexity, the stochastic fluctuation must be at least piecewise continuous in the physical domain \( D \).

   More generally, it could be of interest the construction of a regularity space scale \((X_\delta)_{0 \leq \delta \leq 2}\) interpolating between smooth and rough data corresponding to \( \delta = 0 \) and \( \delta = 2 \) respectively, such that \( r \in X_\delta \) ensures the complexity rate \( e^{-\delta + o(1)} \) for the first/higher order moment algorithms, or for some combination of perturbation and Monte Carlo algorithms (quasi-Monte Carlo).

2. Development of parallel results for the stochastic Galerkin (SG) method which is known to deliver, under stronger assumptions (known joint distributions of the random variables \( (X_m)_{m \in \mathbb{N}_+} \) in the fluctuation expansion), complete statistical informa-
tion (i.e. the probability distribution function) on the stochastic solution $u$.

3. Study of the perturbation algorithm scalability w.r.t. the correlation length of the diffusion coefficient $a$ and the case of a small correlation length, equivalent to a slow (still exponential) convergence of the fluctuation expansion due to unfavourable values of the constants involved in the decay estimates, or to a large pre-asymptotic domain for all results presented in this thesis. This pre-asymptotic analysis should also give some insight into the case of a rough stochastic fluctuation, including white noise, and might require new random field representation techniques, in order to parsimoniously parametrize the data uncertainty, plus an investigation of the applicability of perturbation algorithms in this new context.

4. Generalization and adaptation of perturbation algorithms to the case of unbounded stochastic diffusion coefficient with bounded samples (lower and upper bounds are uniform in the physical domain, but are sample dependent, e.g. log-normally distributed $a$).

5. Extension to general second order elliptic/parabolic/hyperbolic problems with stochastic data (advection-diffusion, linear elasticity, heat, wave equations, Stokes problems etc.).
Chapter 5

Appendix

5.1 A Trace Lemma

**Lemma 5.1.** If $s, t, u$ are real numbers such that

$$t \geq s \geq 0, \quad u \geq s \geq 0 \quad \text{and} \quad t + u - s > d/2,$$

then the mapping (the trace on the diagonal set in $\mathbb{R}^d \times \mathbb{R}^d$)

$$S(\mathbb{R}^{2d}) \ni \phi \rightarrow \Xi(\phi)(x) = \phi(x, x) \in S(\mathbb{R}^d)$$

has a unique linear and continuous extension from $H^t(\mathbb{R}^d) \otimes H^u(\mathbb{R}^d)$ to $H^s(\mathbb{R}^d)$.

**Proof.** Denoting by $F_d$ (and by hat) the Fourier transformation (and transform) in $\mathbb{R}^d$, we compute the Fourier transform of the mapping $\Xi$ (denote by $dx$ the normalized Lebesgue measure in $\mathbb{R}^d$, $dx = (2\pi)^{-d/2} d\lambda$),

$$\hat{\Xi}(\hat{\phi})(\rho) := (F_d \circ \Xi \circ F_d^{-1}(\hat{\phi}))(\rho) = \int_{\mathbb{R}^d} e^{-ix\cdot\rho} (\Xi \circ F_d^{-1}(\hat{\phi}))(x) \, dx$$

$$= \int_{\mathbb{R}^d} e^{-ix\cdot\rho} \int_{\mathbb{R}^{2d}} e^{ix\cdot(\xi + \eta)} \hat{\phi}(\xi, \eta) \, d\xi \, d\eta \, dx$$

$$= \int_{\mathbb{R}^d} e^{-ix\cdot\rho} e^{ix\cdot\theta} \hat{\phi}(\xi, \theta - \xi) \, d\xi \, d\theta \, dx = \int_{\mathbb{R}^d} \hat{\phi}(\xi, \rho - \xi) \, d\xi \quad (5.3)$$

Since the uniqueness of an extension follows by a density argument, we only show that $\hat{\Xi} : S(\mathbb{R}^{2d}) \rightarrow S(\mathbb{R}^d)$ computed in (5.3) has a linear and continuous extension from $K^t(\mathbb{R}^d) \otimes K^u(\mathbb{R}^d)$ to $K^s(\mathbb{R}^d)$ if $s, t, u$ satisfy (5.1) and

$$K^r(\mathbb{R}^d) := L^2(\mathbb{R}^d, \langle x \rangle^{2r} \, dx) \quad \text{with} \quad \langle x \rangle^r := (1 + |x|^2)^{r/2} \quad \forall r \geq 0. \quad (5.4)$$
We estimate therefore
\[ \|\hat{\Xi}(\rho)\|^2_{K^*} = \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \hat{\phi}(\xi, \rho - \xi) d\xi \right)^2 d\rho \]
\[ = \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \frac{\langle \rho \rangle^s}{\langle \xi \rangle^t (\rho - \xi)^{2u}} \cdot \langle \xi \rangle^t (\rho - \xi)^{2u} \cdot \hat{\phi}(\xi, \rho - \xi) d\xi \right)^2 d\rho \]
\[ \leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \frac{\langle \rho \rangle^s}{\langle \xi \rangle^t (\rho - \xi)^{2u}} d\xi \right) \cdot \left( \int_{\mathbb{R}^d} \langle \xi \rangle^t (\rho - \xi)^{2u} \cdot \hat{\phi}(\xi, \rho - \xi)^2 d\xi \right) d\rho \]
\[ \leq \| \int_{\mathbb{R}^d} \frac{\langle \rho \rangle^s}{\langle \xi \rangle^t (\rho - \xi)^{2u}} d\xi \|_{L^\infty(\mathbb{R}^d, d\rho)} \cdot \| \hat{\phi} \|_{K^* \hat{\otimes} K^*}. \quad (5.5) \]

All we have to check is that (5.1) ensures the finiteness of the \( L^\infty \) norm in (5.5). To this end, we note first that it is enough to prove the uniform boundedness of the integral in (5.5) for \( \rho \) outside a compact set in \( \mathbb{R}^d \), e.g. for \( |\rho| > 1 \). Let \( 1/\sqrt{2} < c < 1 \) be an arbitrary real number. We split the integral in two parts, corresponding to the integration domains \( D_1 := \{ \xi \mid \langle \rho - \xi \rangle \geq c \cdot \langle \rho \rangle \} \) and \( D_2 := \{ \xi \mid \langle \rho - \xi \rangle < c \cdot \langle \rho \rangle \} \) and estimate the corresponding integrals as follows
\[ \int_{D_1} \frac{\langle \rho \rangle^s}{\langle \xi \rangle^t (\rho - \xi)^{2u}} d\xi \leq \frac{1}{c^{2s}} \int_{D_1} \frac{1}{\langle \xi \rangle^t (\rho - \xi)^{2(u-s)}} d\xi, \quad (5.6) \]

But it is easy to see that \( \langle \rho - \xi \rangle \geq c \cdot \langle \rho \rangle \) with \( 1/\sqrt{2} < c < 1 \) and \( |\rho| > 1 \) implies \( \langle \rho - \xi \rangle \geq c' \cdot \langle \xi \rangle \) for any \( 0 < c' < \sqrt{2c^2 - 1} \). Inserting in (5.6) we obtain (\( u \geq s \))
\[ \int_{D_1} \frac{\langle \rho \rangle^s}{\langle \xi \rangle^t (\rho - \xi)^{2u}} d\xi \leq \frac{1}{c^{2s} \cdot c^{2(u-s)}} \int_{\mathbb{R}^d} \frac{1}{\langle \xi \rangle^t (\rho - \xi)^{2(u-s)}} d\xi \quad (5.7) \]

and the r.h.s. of (5.7) is finite, due to \( t + u - s > d/2 \), and does not depend on \( \rho \).

To estimate the integral over \( D_2 \), we note first that \( D_2 \) is a ball of center \( \rho \) and radius \( \sqrt{c^2 \cdot \langle \rho \rangle^2 - 1} \), so that \( \langle \xi \rangle \geq c' \cdot \langle \rho \rangle \) with some positive \( c' > 0 \) depending only on \( c < 1 \). It follows
\[ \int_{D_2} \frac{\langle \rho \rangle^s}{\langle \xi \rangle^t (\rho - \xi)^{2u}} d\xi \leq \frac{1}{c^{2s}} \int_{D_2} \frac{\langle \rho \rangle^{2(s-t)}}{\langle \rho - \xi \rangle^{2u}} d\xi \]
\[ \leq \langle \rho \rangle^{2(s-t)} \cdot \int_0^{c(\rho)} r^{2u - d - 1} dr \]
\[ \approx \left\{ \begin{array}{ll}
\langle \rho \rangle^{2(s-t) - 2u + d} & \text{if } u < d/2 \\
\langle \rho \rangle^{2(s-t)} \cdot \log(\rho) & \text{if } u = d/2 \\
\langle \rho \rangle^{2(s-t)} & \text{if } u > d/2 \end{array} \right. \quad (5.8) \]

with constants independent of \( \rho \). The uniform boundedness follows in all cases from the assumptions \( t \geq s \) and \( t + u - s > d/2 \). \( \square \)
5.2 A Fast Decaying Sequence

Here we prove that, if $y, \alpha > 0$ and $j \in \mathbb{N}_+$, the sum of the series with general term $\exp(-y \sum_{k=1}^{j} m_k^\alpha)$ indexed over $1 \leq m_1 < \ldots < m_j < \infty$ is, qualitatively and uniformly in $j \in \mathbb{N}_+$, just as large as the leading term, corresponding to $m_k = k$ for all $1 \leq k \leq j$. More precisely, it holds

**Lemma 5.2.** If $\alpha > 0$, and $x > y > z > 0$, then there exist $c_{\alpha,x,y}, c_{\alpha,y,z} > 0$ such that

$$c_{\alpha,x,y} \exp\left(-x \frac{1}{1+\alpha} j^{1+\alpha}\right) \leq \sum_{1 \leq m_1 < \ldots < m_j < \infty} \prod_{k=1}^{j} \exp(-ym_k^\alpha) \leq c_{\alpha,y,z} \exp\left(-z \frac{1}{1+\alpha} j^{1+\alpha}\right) \quad (5.9)$$

$\forall j \in \mathbb{N}_+$.

**Proof.** For $y > 0$ and $j, J \in \mathbb{N}_+ \cup \{\infty\}$ with $j \leq J$ we set

$$S_{j,J} := \sum_{1 \leq m_1 < \ldots < m_j \leq J} \prod_{k=1}^{j} \exp(-ym_k^\alpha). \quad (5.10)$$

We check first the lower bound in (5.9). To this end we note that the sum in (5.10) contains the term corresponding to $m_k = k$ for all $1 \leq k \leq j$, therefore we have

$$S_{j,J} \geq \exp(-y \sum_{k=1}^{j} k^\alpha). \quad (5.11)$$

The conclusion follows by noting that

$$\sum_{k=1}^{j} k^\alpha \leq (j+1)^{1+\alpha} \int_{0}^{1} x^\alpha dx = (j+1)^{1+\alpha} \frac{1}{1+\alpha}. \quad (5.12)$$

It remains to prove the upper bound of the sum in (5.9). From the definition (5.10) it follows that the sequence $(S_{j,j})_{j \in \mathbb{N}_+}$ is rapidly decaying, that is

$$S_{j,j} \leq c_{\alpha,y,z}\beta^j \quad \forall j \in \mathbb{N}_+, \forall \beta > 0. \quad (5.13)$$

From (5.10) we derive the recursive formula

$$S_{j,J+1} = S_{j,J} + \exp(-y(J+1)^\alpha) \sum_{1 \leq m_1 < \ldots < m_j \leq J} \prod_{k=1}^{j-1} \exp(-ym_k^\alpha) = S_{j,J} + \exp(-y(J+1)^\alpha)S_{j-1,J} \quad (5.14)$$

From (5.14) it follows immediately by induction on $j$ that

$$S_{j,J} < S_{j,\infty} = \lim_{J \to \infty} S_{j,J} < \infty \quad \forall j \in \mathbb{N}_+, \quad (5.15)$$
and that
\[ S_{j,\infty} \leq S_{j,j} + \sum_{k=j+1}^{\infty} \exp(-yk^\alpha) \cdot S_{j-1,\infty}. \] (5.16)

Now, for an arbitrary \( \gamma \in (0, 1) \) we have, for \( j \) large \( (j \geq j_0, \) with \( j_0 \) depending on \( y, \alpha, \gamma)\),
\[ \sum_{k=j+1}^{\infty} \exp(-yk^\alpha) \leq \gamma, \] (5.17)
which ensures
\[ S_{j,\infty} \leq S_{j,j} + \gamma S_{j-1,\infty}, \quad \forall j \geq j_0. \] (5.18)

From (5.13) and (5.18) we deduce that
\[ S_{j,\infty} \leq c_{\alpha, \gamma, \beta}((\gamma + \beta)^j + \gamma^{j-j_0+1} S_{j_0-1,\infty}), \] (5.19)
which shows that \( S_{j,\infty} \to 0 \) as \( j \to \infty \), by choosing \( \beta \) such that \( \gamma + \beta < 1 \). The sequence \((S_{j,\infty})_{j \in \mathbb{N}_+}\) is in particular bounded, that is
\[ S_{j,\infty} \leq c_{\alpha, \gamma} \quad \forall j \in \mathbb{N}_+. \] (5.20)

Since this inequality holds for any \( y > 0 \), the conclusion follows then from (5.20) upon replacing \( y \) by \( y - z \). \( \square \)
Bibliography


Curriculum Vitae

Name: Radu Alexandru Todor
Date of Birth: November 12, 1973
Place of Birth: Brasov, Romania
Nationality: Romanian

Education
1980-1992
Primary School/Secondary School/College
"Andrei Saguna" Brasov, Romania
1988, '89, '90
First Prize, Romanian National Mathematical Olympiad
1991
Bronze Medal, International Mathematical Olympiad, Sigtuna, Sweden
1992
Gold Medal, Jury Prize, Balkan Mathematical Olympiad, Athens, Greece
Silver Medal, International Mathematical Olympiad, Moscow, Russia
1992-1997
Study of Mathematics, Bucharest University
B. Sc. Diploma Thesis: Morse Inequalities on Covering Manifolds
Advisor: Prof. V. Iftimie
1997-1999
Post Graduate Studies, Bucharest University
Master Program: Nonlinear Equations and Modelling
Master Thesis: Spectral Theory for Uniformly Propagative Systems
Advisor: Prof. V. Iftimie
1999
(April-June) Research Stay at Department of Mathematics,
Université Cergy-Pontoise, Paris, France
2000-present
Ph. D. Student and Research Assistant, ETH Zurich
Advisor: Prof. Ch. Schwab