

# Multilevel Monte Carlo Methods for Uncertainty Quantification in Brain Simulations



Matteo Croci  
University College  
University of Oxford

A thesis submitted for the degree of  
*Doctor of Philosophy*  
Long Vacation 2019

*A Tiziano e Clara.  
Al Nin e a Virginia.  
A Pierluigi e a Luisa.*

## Acknowledgements

First and foremost, I would like to express my deepest thanks to my amazing supervising trio: Professor Patrick E. Farrell, Professor Michael B. Giles and Professor Marie E. Rognes, for all their support and guidance throughout the process of writing this thesis. Without their vision, clarity of thought, passion and excitement for the subject I would have never got this far. It has truly been a pleasure and an honour to work with them and I am deeply thankful for everything they have taught me. Informally, it has been a lot of fun.

I would also like to thank Professor Endre Süli, Professor Christoph Reisinger, Dr. Abdul-Lateef Haji-Ali and Dr. Alberto Paganini for being my examiners for my transfer and confirmation of status examinations and for all the valuable comments and suggestions. I would also like to thank Professor Endre Süli and Professor Robert Scheichl for accepting to be my examiners for my viva.

Throughout my DPhil, I have had the privilege of receiving help from and discussing about my work with many brilliant and talented researchers. For this reason, I would like to thank Professor Endre Süli (again), Professor Andrew Wathen, Dr. Abdul-Lateef Haji-Ali, Dr. Alberto Paganini, Dr. James Maddison, Dr. Vegard Vinje, and (soon to be Dr.) Casper Beentjes for being so generous with their time.

The numerical experiments in my thesis required quite some computational power. I am deeply grateful to the Oxford Mathematical Institute IT staff and to the Department of Research Computing of the University of Oslo IT department for all the assistance throughout the years. In this regard, I would also like to express my gratitude to Johannes Ring and Martin Sandve Alnæs for all their help in the initial and final stages of my DPhil.

A big thanks also goes to Simula Research Laboratory for co-funding my DPhil. To all the people and employees of Simula, thank you so much for all the hospitality and support, for all the astonishingly fun trips and for making me feel at home every time I come to Oslo.

I would furthermore like to thank:

- The InFoMM team for setting up and running such an amazing DPhil programme.
- The Oxford Numerical Analysis group for the brilliant academic research environment.
- The EPSRC, University College and the Oxford-Radcliffe scholarship anonymous donors, for the financial support without which this thesis would not have been written.

- University College and St. Anne's College for providing me with such an amazing social environment through which I have met so many wonderful people.
- My officemates Florian Wechsung and Federico Danieli for all the fun, the songs and for putting up with me throughout the years.
- The University of Oxford Triathlon club for overlooking my debatable triathlon skills and still accepting me as a member of the team.
- The University of Oxford Public International Law students for *de facto* offering me honorary membership of their group without ever having studied law. They still think this thesis is about triangles. A special thanks goes to Daniel Franchini, the "Greek gang" and associates for all the fun (academic and not) in the last 4-5 years.
- The unbeatable first year St. Anne's crew for staying in touch even if scattered around the globe.
- My Oxford Italian and Spanish friends for their friendship, for all the nights out and for the delicious dinners. Thank you for keeping me sane and making me feel at home in the last few years.
- All my Italian friends from back home thanks for still managing to stay in touch after all these years. It is a privilege to be your friend.
- Virginia Vélaz Azcoiti for having always been by my side, for her special ingredient-free recipes, and for her unconditional support.
- My parents, grandparents and family for all the unwavering support they have given me since I was born.

## **Statement of Originality**

This thesis is entirely my own work except where otherwise indicated. I have clearly signalled the presence of quoted or paraphrased material and referenced all sources. Except for Chapter 6, all chapters have not been submitted, either wholly or substantially, for another Honour School or degree of this University, or for a degree at any other institution. Parts of Chapter 6 were drawn from the paper [44], which was jointly written with Vegard Vinje. Vinje submitted this paper as part of his PhD at Simula Research Laboratory/the University of Oslo. Further details about this collaboration are written at the beginning of Chapter 6.

# Abstract

This thesis consists of two parts. In the first, we develop two new strategies for spatial white noise and Gaussian-Matérn field sampling that work within a non-nested multilevel (quasi) Monte Carlo (ML(Q)MC) hierarchy. In the second, we apply the techniques developed to quantify the level of uncertainty in a new stochastic model for tracer transport in the brain.

The new sampling techniques are based on the stochastic partial differential equation (SPDE) approach, which recasts the sampling problem as the solution of an elliptic equation driven by spatial white noise. We present a new proof of an *a priori* error estimate for the finite element (FEM) solution of the white noise SPDE. The proof does not require the approximation of white noise in practice, and includes higher order elliptic operators and  $p$ -refinement.

Within the SPDE approach, the efficient sampling of white noise realisations can be computationally expensive. In this thesis, we present two new sampling techniques that can be used to efficiently compute white noise samples in a FEM-MLMC and FEM-MLQMC setting. The key idea is to exploit the finite element matrix assembly procedure and factorise each local mass matrix independently, hence avoiding the factorisation of a large matrix. In a multilevel framework, the white noise samples must be coupled between subsequent levels. We show how our technique can be used to enforce this coupling even in the case of non-nested mesh hierarchies.

In the MLQMC case, the QMC integrand variables must also be ordered in order of decaying importance to achieve fast convergence with respect to the number of samples. We express white noise as a Haar wavelet series whose hierarchical structure naturally exposes the leading order dimensions. We split this series in two terms which we sample via a hybrid standard Monte Carlo/QMC approach.

We demonstrate the efficacy of our sampling methods with numerical experiments. In a multilevel setting, a good coupling is enforced and the telescoping sum is respected. In the MLQMC case, the asymptotic convergence rate is the same as standard Monte Carlo, but significant computational gains are obtained in practice thanks to a pre-asymptotic QMC-like regime.

In the final part of the thesis, we employ a combination of the methods presented to solve a PDE with random coefficients describing tracer transport within the interstitial fluid of the brain. Numerical simulations support the claim that diffusion alone cannot explain the penetration of tracers within deep brain regions as observed in clinical experiments, even when uncertainties in the diffusivity have been accounted for. A convective velocity field may however increase tracer transport, provided that a directional structure is present in the interstitial fluid circulation.

# Contents

<b>1</b>	<b>Introduction and physiological background</b>	<b>1</b>
1.1	DPhil thesis overview . . . . .	1
1.2	Introduction . . . . .	1
1.2.1	CDT partner company: Simula . . . . .	1
1.2.2	Project motivation: brain disease . . . . .	1
1.2.3	Model uncertainty and uncertainty quantification . . . . .	3
1.2.4	Thesis outline . . . . .	3
1.3	Physiological background . . . . .	4
1.3.1	The fluid dynamics of the brain . . . . .	4
1.3.2	The glymphatic hypothesis and related theories in the mathematical and biomedical computing literature . . . . .	5
<b>2</b>	<b>Mathematical background</b>	<b>8</b>
2.1	Monte Carlo methods . . . . .	8
2.1.1	Standard Monte Carlo . . . . .	8
2.1.2	Randomised quasi Monte Carlo . . . . .	10
2.2	Multilevel Monte Carlo methods . . . . .	14
2.2.1	Multilevel Monte Carlo . . . . .	14
2.2.2	Multilevel quasi Monte Carlo . . . . .	17
2.3	Standard and generalised stochastic fields . . . . .	18
2.4	Gaussian field sampling . . . . .	21
2.4.1	Matrix factorisations . . . . .	21
2.4.2	Karhunen-Loëve expansion . . . . .	23
2.4.3	Circulant embedding . . . . .	24
2.4.4	The SPDE approach to Matérn field sampling . . . . .	26
<b>3</b>	<b>Finite element convergence analysis for spatial white noise driven elliptic PDEs</b>	<b>29</b>
3.1	Finite element convergence for problem (2.43): the case $k = 1$ . . . . .	30
3.1.1	Assumptions on the domain $D$ . . . . .	31
3.1.2	Known techniques: introducing an auxiliary problem . . . . .	32
3.1.3	The main result . . . . .	33
3.2	Proof of Theorem 3.2 . . . . .	34
3.2.1	Our “approximation” of white noise . . . . .	34
3.2.2	Main part of the proof . . . . .	38
3.3	Finite element convergence for problem (2.43): the case $k > 1$ . . . . .	41

<b>4 Efficient white noise sampling and coupling for MLMC</b>	<b>45</b>
4.1 Background and existing literature . . . . .	45
4.2 The white noise sampling problem . . . . .	47
4.2.1 Finite element solution of elliptic PDEs with white noise forcing . . . . .	47
4.2.2 Multilevel white noise sampling/white noise coupling condition . . . . .	48
4.2.3 Embedded meshes and non-nested grids . . . . .	50
4.3 Supermesh constructions between quasi-uniform meshes . . . . .	50
4.4 White noise sampling . . . . .	55
4.4.1 Sampling of independent white noise realisations . . . . .	56
4.4.2 Sampling coupled white noise realisations for MLMC . . . . .	57
4.5 Numerical results . . . . .	62
4.5.1 Matérn field convergence . . . . .	64
4.5.2 MLMC convergence . . . . .	69
4.6 Chapter conclusions . . . . .	73
<b>5 Multilevel quasi Monte Carlo methods for white noise driven elliptic PDEs</b>	<b>74</b>
5.1 Related work . . . . .	74
5.2 Haar wavelet expansion of spatial white noise . . . . .	77
5.3 Sampling independent realisations for QMC . . . . .	79
5.3.1 Sampling of $\dot{W}_{\mathcal{L}}$ . . . . .	81
5.3.2 Sampling of $\dot{W}_R$ . . . . .	85
5.4 A partial QMC convergence result . . . . .	90
5.5 Sampling coupled realisations for MLQMC . . . . .	94
5.6 Numerical results . . . . .	98
5.6.1 Testing the supermesh construction . . . . .	98
5.6.2 Testing the sampling strategies . . . . .	102
5.6.3 MLQMC convergence . . . . .	103
5.7 Chapter summary and conclusions . . . . .	110
<b>6 Uncertainty quantification of tracer distribution in the brain using random interstitial fluid velocity fields</b>	<b>111</b>
6.1 Introduction . . . . .	111
6.2 Stochastic modelling of brain tracer concentration . . . . .	114
6.2.1 The baseline model . . . . .	114
6.2.2 Stochastic diffusion modelling . . . . .	118
6.2.3 Stochastic velocity modelling . . . . .	119
6.3 Numerical solution of the stochastic models . . . . .	124
6.4 Clinical results: standard Monte Carlo simulations . . . . .	128
6.4.1 Numerical results . . . . .	128
6.4.2 Discussion . . . . .	136
6.4.3 Section conclusions . . . . .	138
6.5 Computational comparison of the methods presented in this thesis . . . . .	138
6.5.1 Numerical results . . . . .	138
6.5.2 Section conclusions . . . . .	143
<b>7 Conclusions and future research directions</b>	<b>144</b>
7.1 Summary of results . . . . .	144
7.2 Future research directions . . . . .	145

# Chapter 1

## Introduction and physiological background

### 1.1 DPhil thesis overview

### 1.2 Introduction

#### 1.2.1 CDT partner company: Simula

This thesis is based on work supported by the EPSRC Centre For Doctoral Training in Industrially Focused Mathematical Modelling (EP/L015803/1) in collaboration with Simula Research Laboratory.

Simula is a research laboratory with the main objective of creating knowledge about fundamental scientific challenges that are of genuine value for society. The mission of Simula is to perform important research with long-term impact in the fields of communication systems, scientific computing and software engineering. The strong focus on research is combined with both teaching of postgraduate students and the development of commercial applications<sup>1</sup>.

The Department of Numerical Analysis and Scientific Computing (SCAN) in Simula is one of the driving forces behind the FEniCS Project [6, 120]. Within SCAN, Marie E. Rognes' group focuses on the simulation of physiological processes that affect human health. The expertise of the group ranges from medical research and bioengineering to mathematical modelling, numerical analysis and scientific computing. One of the group's current interests is the simulation of the fluid dynamics within the brain, which in this thesis we refer to as the brain 'waterscape'.

#### 1.2.2 Project motivation: brain disease

The motivation for our project is the analysis of the symptoms and causes that can lead to brain diseases. In fact, anomalies in the brain waterscape are believed to

---

<sup>1</sup>Adapted from [www.simula.no/about-simula-research-laboratory](http://www.simula.no/about-simula-research-laboratory).

be related to common diseases such as Alzheimer's and other forms of dementia [117, 156, 174, 181].

Dementia is a generic term that describes a wide range of diseases which have the common denominator of creating a disorder in a person's mental processes. Examples can be memory loss, personality changes, impaired reasoning and thinking, deterioration of social behaviour, motivation or emotional control [178]. Worldwide, about 47 million people suffer from dementia and every year 7.7 millions more cases are diagnosed, with an expected 75.6 million people being affected by 2030. The cost of dementia health care is enormous. This amounts to almost \$604 billion per year in the US and around £25 billion per year in the UK, a value which is constantly increasing [7, 178]. '*More research is needed to develop new and more effective treatments and to better understand the causes of dementia*' (World Health Organization [178]).

The most well-known type of dementia is Alzheimer's disease. One of the hallmark characteristics of Alzheimer's is the formation of plaques made up of a protein called Amyloid- $\beta$ . This protein can be normally found as a solute in the cellular cytosol and in the brain interstitial fluid (ISF) (present in the brain extracellular space, the space between brain cells) [124]. However, under Alzheimer's conditions, this protein accumulates in the brain forming plaques that may induce neuronal death.

Every cell in our body absorbs its nutrients from the blood and then releases its metabolic wastes in the surrounding extracellular space. From there, the metabolic waste is cleared by the lymphatic system, a network of tissues and organs specialised in toxin and waste clearance. The only part of our body where (almost) no lymphatic vessels are present is the brain, where some other unknown clearance mechanism is in action. One of the leading hypotheses in Alzheimer's research is that when this mechanism, whatever it might be, is impaired, Amyloid- $\beta$  accumulates damaging the brain and causing Alzheimer's. If researchers could discover what this mechanism is, they might possibly prevent the damage and find a cure.

For this reason, over the last few years the study of the physiological mechanisms governing the movement of fluids in the brain has gained prominence [96, 146, 147, 179]: understanding how the brain waterscape works can help discover how dementia develops. However, experimenting with the human brain *in vivo* is extremely difficult and the subject is still poorly understood. This makes this topic a highly active interdisciplinary research area. Invasive experiments are performed on mice, but their results are not necessarily representative of the human physiology. In addition, they cannot be performed on people as that would be unethical. Non-invasive experiments, instead, do not give us enough information and the measurement itself can affect the physics of what is being observed [53]. For this reason, *in silico* simulations of these phenomena are gaining popularity [9, 26, 48, 91, 92, 97, 118, 161] as they represent an alternative avenue of investigation that is cheap and does not require human experimentation.

### 1.2.3 Model uncertainty and uncertainty quantification

Numerical simulations based on mathematical models rely on physiological information that is obtained through *in vivo* measurements. Modern techniques allow the measurement of most physical parameters with a reasonable degree of accuracy [8, 64, 72, 164]. Nevertheless, measurement errors are still present [77, 94]. This is often the source of discordant opinions in the literature [64, 117, 153, 154]. One example is given by the brain tissue permeability which varies up to 7 orders of magnitude in computational models [79, 92].

Overall, one of the main challenges in brain simulation is the lack of accurate quantitative information on the mechanical input parameters needed to set up mathematical models. Quantities such as brain matter permeability, interstitial fluid flow velocity and diffusivity are only known approximately or on average. The position of the blood vessels and capillaries can be measured, but it varies from patient to patient and it is extremely difficult to resolve without significant expense.

The focus of this thesis is estimating how uncertainty in these input parameters affects models and simulations' predictions and propagates to output quantities of clinical interest. Quantifying this uncertainty is extremely important for medical applications. We construct surrogates for the physiological parameters that account for the uncertainty in their values through the use of Gaussian random fields and we use multilevel Monte Carlo methods to perform forward uncertainty quantification (UQ) on brain fluid models. We solve the mathematical problems of interest via the finite element method (FEM).

### 1.2.4 Thesis outline

The main content of this thesis can be logically divided in two parts. In the former (chapters 3, 4 and 5) we focus on a family of numerical methods for Gaussian-Matérn field sampling and we present a new sampling strategy to use within the framework of multilevel Monte Carlo methods. In Chapter 5, we also develop a new multilevel quasi Monte Carlo method for partial differential equations (PDEs) with random coefficients. In the latter part of the thesis (Chapter 6), we apply these methods to quantify the uncertainty in a new model for brain solute concentration and fluid flow through the poroelastic brain matter.

We begin the thesis with a short overview of the physiological background needed (Section 1.3). In Chapter 2, we describe the mathematical background needed for the rest of the thesis, including a review of Monte Carlo methods and of the existing techniques for Gaussian field sampling. Sampling Gaussian field realisations can be extremely expensive computationally and various techniques are available [49, 62, 84, 116]. The method we focus on in our work is referred to as the *SPDE approach to Matérn field sampling* [116] and it consists of recasting the Gaussian field sampling problem as the solution of an elliptic PDE driven by spatial white noise. In Chapter 3 we derive *a priori* FEM convergence estimates for the solution of spatial white noise driven elliptic PDEs. In Chapter 4 we present a novel technique for efficient white noise sampling for MLMC with non-nested meshes. In Chapter 5 we extend

this technique to the multilevel quasi Monte Carlo case. In Chapter 6 we consider a convection-diffusion equation with random velocity and diffusivity fields for the concentration of a tracer soluted in the interstitial fluid of human brain, which we solve with the new techniques presented in the thesis. We conclude the thesis with Chapter 7.

## 1.3 Physiological background

### 1.3.1 The fluid dynamics of the brain

The cerebral matter, or parenchyma, is comprised of two types of tissues: the gray matter and the white matter. The former is mostly composed by neuron cell bodies, dendrites, myelinated and unmyelinated axons, and glial cells, while the latter is mostly comprised of myelinated cell axons connecting together various gray matter regions. The gray matter surrounds the white matter in the brain (see Figure 1.1). As reported in [72], gray and white matter have different mechanical properties. In particular, white matter is stiffer than gray matter and presents shear anisotropy, larger regional variation, and increased viscosity.

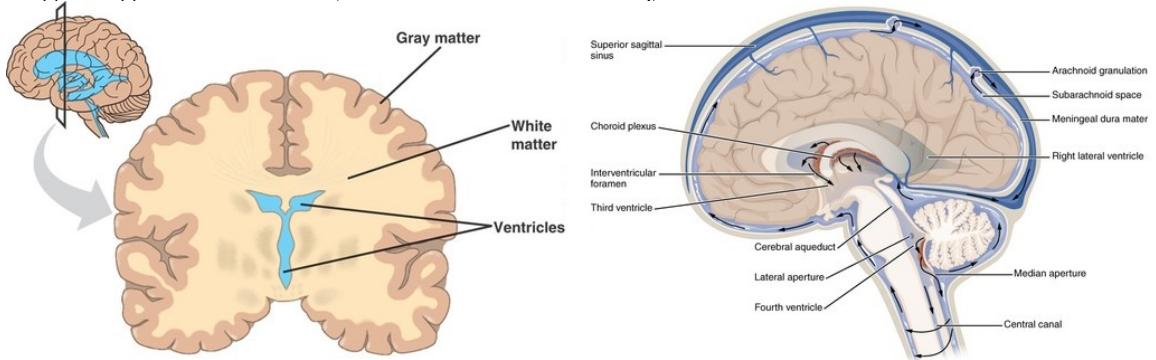


Figure 1.1: *Brain charts: white and gray matter (left), ventricles (left, light blue) and CSF flow (right, image taken from OpenStax, Anatomy & Physiology). The ISF flows within the parenchyma.*

The human brain is influenced by three different types of flow networks: the blood vasculature, the cerebrospinal fluid (CSF) (Figure 1.1) and the interstitial fluid (ISF). These networks are separated by semi-permeable membranes, namely the brain-blood barrier (ISF-blood) and the blood-CSF barrier, which selectively only allow exchange of certain substances and liquids.

The vascular network comprises the arterial, venous, and capillary blood flow. It carries the nutrients to the brain cells within the parenchyma (the cerebral matter) and it is strongly related to the production of CSF, which is believed to be produced from blood plasma in the choroid plexuses [117]. Each choroid plexus is situated within one of the four brain ventricles, which are filled by CSF (Figure 1.1). From the ventricles, the CSF flows in the spinal canal and in the subarachnoid space, the space between the arachnoid and pia meninges. One hypothesis is that the CSF is reabsorbed in the subarachnoid space by the arachnoid granulations [101] (Figure 1.1).

Brain cells absorb their nutrients from the vascular system and release metabolic waste into the interstitial space, the space between the cells. Here, the waste is absorbed and collected by the ISF, which permeates the brain tissue [117]. The interstitial space and the ISF make the parenchyma a sponge-like solid with poroviscoelastic properties and permeated by liquid [72].

The dynamics of ISF flow has been extensively studied with the use of molecular tracers [1], but it is not fully understood [1, 117]. One hypothesis is that the ISF produced in the brain tissue can be reabsorbed by the capillary system so as to regulate brain water content. This hypothesis has been used in the mathematical literature [171]. It is also uncertain whether ISF can flow from the parenchyma to the subarachnoid space [117].

ISF flow is believed to be driven by deformations of the arachnoid interface due to pressure oscillations in the cerebral vasculature [117]. At every heart beat, blood is pumped in the brain vasculature, causing a pressure rise and a volume dilatation during systole. The pressure drops, and the brain volume contracts during diastole. Whether this volume variation is due to volume variation in the brain tissue or in the brain surface arteries is still debated.

### 1.3.2 The glymphatic hypothesis and related theories in the mathematical and biomedical computing literature

One of the main hypothesis about brain metabolic waste clearance is called the glymphatic hypothesis, according to which a bulk flow (which is a consequence of advection-dominated fluid movement) of ISF has been proposed to occur through the brain parenchyma [96]. This directional fluid movement has been named the glymphatic system, with bulk flow being a mechanism for effective waste clearance from the brain parenchyma. The glymphatic hypothesis is strongly dependent on the assumption that a strong bulk flow is present within the brain and is still far from being established. In fact, some aspects of ISF movement are still not well understood. The exchange between CSF and ISF are believed to occur along small fluid filled spaces surrounding large penetrating arteries in the brain parenchyma known as paravascular spaces (PVS) [93, 145], cf. Figure 1.2. However, while evidence of influx of tracer along paravascular spaces seems evident, the outflux route is more debated [37].

Discovering how the brain clears itself from metabolic waste could be one of the milestones to reach before being able to understand how Alzheimer's disease develops. For this reason, several attempts have been made in the mathematical and biomedical computing literature to reproduce, verify or disprove the glymphatic hypothesis [9, 26, 48, 92, 97, 118, 161]. In Chapter 6 of this thesis we also consider the glymphatic hypothesis: we take from the clinical experiments of Ringstad et al. [146] and investigate whether their experimental results can be reproduced *in silico* under the glymphatic hypothesis, while accounting for the uncertainty in the input parameters of our model. We therefore conclude this chapter with a brief literature review about the mathematical and biomedical computing studies related to brain solute movement

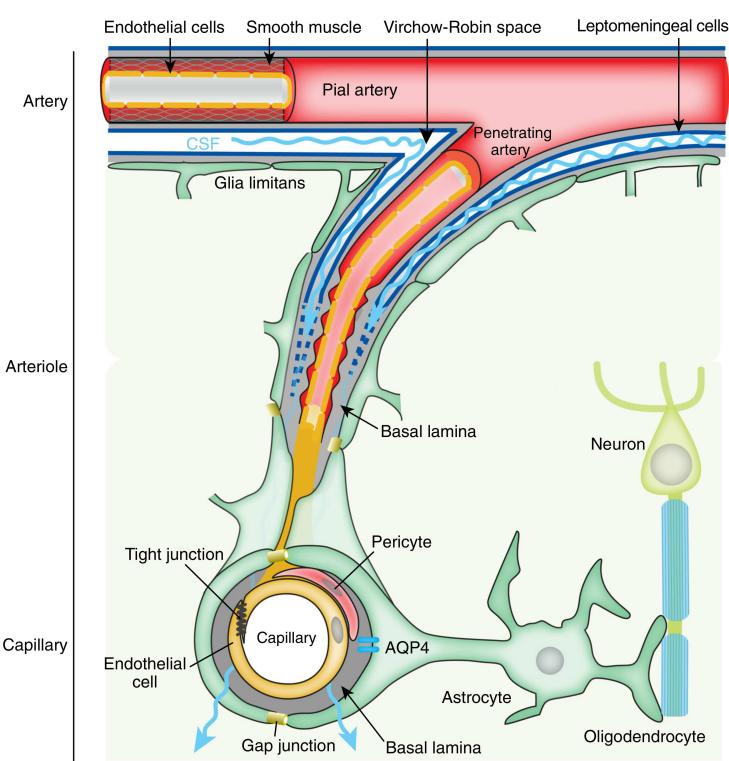


Figure 1.2: *A schematic of paravascular CSF circulation. Exchange between CSF and ISF has been hypothesised to happen along paravascular spaces (also called Virchow-Robin spaces).* Reprinted by permission from RightsLink/Springer Nature: *The Glymphatic System: A Beginner’s Guide* by Jessen et al., Neurochemical Research, copyright ©2015, Springer Science & Business Media, New York [96].

and/or the glymphatic hypothesis. For a more general review about the mathematical modelling of brain mechanics, we refer to the excellent review article by Goriely et al. [72].

A first class of numerical studies investigates the movement of CSF along the PVS and possibly its relation to solute drainage. In [26], Bilston et al. research whether arterial pulsations are sufficient to drive CSF flow along the PVS in the spinal cord. They model the PVS as a radially symmetric annulus of which one boundary, the blood vessel wall, deforms according to a train of pulse waves representing arterial pulsations. They model PVS flow with the Navier-Stokes equation in a low Reynolds number regime which they solve using a finite volume method on a moving mesh.

While it is now established that indeed vessel pulsations have an active role in the CSF influx into the parenchyma, there is evidence suggesting that solutes do not leave the parenchyma following the same route [37]. One of the leading alternative hypothesis, the intramural periarterial drainage theory (IPAD), states that solutes are removed from the brain along the basement membranes of capillaries and arteries. The exact mechanism is, however, still under investigation [37]. In [48] Diem et al. investigate whether arterial pulsations alone can be the driving force of IPAD and conclude that a valve mechanism (e.g. directional permeability) is required to achieve enough net reverse flow to justify the drainage. Diem et al. use Darcy’s law with a

pressure-dependent permeability and lubrication theory to derive a one dimensional IPAD model in cylindrical coordinates, which they solve using a finite volume method.

A second class of studies directly tackles the main assumption underlying the glymphatic hypothesis: the presence of a strong bulk flow of ISF across the parenchyma transporting solutes by advection. All these numerical studies seem to conclude that advection alone is not sufficient to explain the movement of tracer and solutes observed in magnetic resonance imaging (MRI) studies, but rather that diffusion [92, 97] or a combination of both advection and diffusion [9] is the leading mechanism. In [97], Jin et al. solve the Navier-Stokes and convection-diffusion equations using a realistic ECS geometry in 2D to investigate short-range transport of solutes and they conclude that a strong bulk flow requires a sustained high pressure difference, which cannot be caused by arterial pulsations. Introducing diffusion is however sufficient to match experimental studies. Asgari et al. in [9] use the same equations on a 3D axisymmetric channel with a pulsating boundary representing arterial pulsations and a dispersion equation in a 1D domain modelling solute movement between paraarterial and paravenous spaces. Their results indicate that arterial pulsations are unlikely to be driving solute transport and that instead dispersion could be responsible. Holter et al. in [92] use Stokes and the convection-diffusion equations in a real-life 3D geometry of the ECS which they solve with the FEM. Holter et al. discover that the ECS permeability is too low to allow strong bulk flow and that diffusion is more likely to be the primary cause of solute transport.

The third and last class of studies does not only focus on fluid movement, but also simulates the movement of solutes within the ISF and the process of drug delivery. Linninger et al. in [118] use MRI data to construct a 3D brain mesh on which they simulate the effect of intraparenchymal drug delivery through a catheter. They use the convection-diffusion equation in the catheter and across the brain combined with the Navier-Stokes equations in the catheter and porous flow equations within the brain tissue. Støverud et a. in [161] solve a diffusion tensor imaging assisted porous flow model on a real-life 3D brain geometry and study how drug transport is affected by the fact that ISF flows within a poro-elastic deformable media.

In Chapter 6 we study the effect of random spatial variations in the velocity and diffusivity fields due to patient variability, measurement errors and/or lack of knowledge. Additionally, we test different assumptions about ISF fluid movement and solute clearance, including the glymphatic hypothesis and the IPAD. Until then we are only going to consider the numerical techniques for UQ needed for this investigation.

# Chapter 2

## Mathematical background

Throughout the thesis, the main objective will be to derive Monte Carlo sampling methods for the solution of PDEs with random field coefficients that model the movement of tracers and solutes within the brain interstitial fluid. These PDEs will all take the form: find the ISF solute/tracer concentration  $c(t, \mathbf{x}, \omega)$  such that,

$$\dot{c} + \nabla \cdot (\mathbf{v}(\mathbf{x}, \omega)c) - \nabla \cdot (D^*(\mathbf{x}, \omega)\nabla c) + rc = f(\mathbf{x}), \quad \omega \in \Omega, \quad \mathbf{x} \in G, \quad t > 0, \quad (2.1)$$

with suitable boundary conditions to be discussed in Chapter 6. Here  $\Omega$  is a sample space and  $G \subseteq \mathbb{R}^3$  is the parenchyma domain. The coefficients  $\mathbf{v}(\mathbf{x}, \omega)$  and  $D^*(\mathbf{x}, \omega)$  are stochastic fields (what stochastic fields are will be defined later in this chapter) and the solution  $c(t, \mathbf{x}, \omega)$  is itself random. In this thesis we consider the estimation of the expectation of an output functional  $\mathcal{P}(c) \equiv P(\omega)$  of the solution. We refer to Chapter 6 for further details, but it is useful to keep in mind that solving this problem requires the efficient sampling of  $\mathbf{v}(\mathbf{x}, \omega)$  and  $D^*(\mathbf{x}, \omega)$  from their distribution, the numerical solution of (2.1) via the finite element method, and the estimation of  $\mathbb{E}[P]$  via Monte Carlo methods.

In this chapter we give an overview of the main theoretical and computational ingredients needed in the rest of the thesis. First, we give an overview of the Monte Carlo (MC), quasi Monte Carlo (QMC), multilevel Monte Carlo (MLMC) and multilevel quasi Monte Carlo (MLQMC) methods. We then introduce the concepts of stochastic fields and generalised stochastic fields. Among these, of central importance in this work are Matérn-Gaussian fields and Gaussian white noise. The last section in this chapter gives a highlight of existing techniques for Gaussian field sampling and describes in detail the SPDE approach to Matérn field sampling that we use throughout the thesis.

### 2.1 Monte Carlo methods

#### 2.1.1 Standard Monte Carlo

For a given sample space  $\Omega$ ,  $\sigma$ -algebra  $\mathcal{A}$  and probability measure  $\mathbb{P}$ , let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space and let  $L^2(\Omega, \mathbb{R})$  indicate the space of scalar real-valued random

variables with finite second moment. The Monte Carlo method (MC) is a numerical technique used to estimate the expected value of a random variable of interest  $P(\omega) \in L^2(\Omega, \mathbb{R})$ ,  $\mathbb{E}[P]$ . Given  $N$  independent samples  $P(\omega^n)$ ,  $n = 1, \dots, N$ , the Monte Carlo estimator reads,

$$\mathbb{E}[P] \approx \hat{P} = \frac{1}{N} \sum_{n=1}^N P(\omega^n). \quad (2.2)$$

It is important to know how large  $N$  should be to achieve a given error tolerance. The error measure we adopt is the mean square error (MSE), namely,

$$MSE = \mathbb{E}[(\hat{P} - \mathbb{E}[P])^2]. \quad (2.3)$$

This can be rewritten in a more insightful way [68] as

$$MSE = \mathbb{V}[\hat{P}] + \mathbb{E}[P - \hat{P}]^2 = \frac{\bar{V}}{N} + \mathbb{E}[P - \hat{P}]^2, \quad (2.4)$$

where  $\bar{V}$  is the variance of a single Monte Carlo sample  $P(\omega^n)$  and the quantity  $|\mathbb{E}[P - \hat{P}]|$  is called the weak error. The first term on the RHS depends on the variance of the approximation, which is related to the actual variance of  $P$ . This term can be decreased by increasing the number of samples  $N$ . The second term on the RHS is, instead, related to the error made in the numerical approximation of  $P$ : if  $P(\omega^n)$  can be sampled exactly then the weak error is zero since  $\mathbb{E}[P] = \mathbb{E}[\hat{P}]$ . In practical applications, e.g. when  $P$  is the output functional of a PDE with random coefficients, computing  $P$  exactly is not an option and the weak error must be controlled as well.

Say that we want to achieve a root MSE of  $O(\varepsilon^2)$ , then we would need both variance and weak error terms to be of  $O(\varepsilon^2)$ . Assuming that the computational cost required to sample  $P$  accurately enough to have an  $O(\varepsilon)$  weak error is  $O(\varepsilon^{-q})$ , for  $MSE = O(\varepsilon^2)$  we need  $N = O(\bar{V}\varepsilon^{-2})$  samples for a total MC cost complexity of  $\bar{C} = O(\bar{V}\varepsilon^{-2-q})$ .

The Monte Carlo method has four big advantages: 1) provided that  $P \in L^2(\Omega, \mathbb{R})$  no further assumptions are required on the smoothness of  $P$  with respect to  $\omega$ , 2) standard MC does not suffer from the curse of dimensionality when approximating the expectation of random variables with large stochastic input dimensions, 3) each sample in the MC estimator is independent and can therefore be computed in parallel, 4) the variance of  $P$  can be estimated by MC concurrently with the expectation providing an approximate confidence interval for the MC approximation. However, this comes at a price: the slow  $O(N^{-1/2})$  Monte Carlo convergence rate results in a large  $O(\bar{V}\varepsilon^{-2-q})$  computational cost. Let us fix our ideas with an example: assume that we require  $\varepsilon^2 = 0.01^2$  MSE accuracy and that the variance of  $P$  is  $O(1)$ ; we then need approximately  $10^4$  samples of  $P$  of sufficient accuracy. If the sampling of  $P$  requires the solution of a PDE (this is the case in this thesis), i.e.  $q$  is large, we would then need approximately  $10^4$  PDE solves for convergence.

### 2.1.2 Randomised quasi Monte Carlo

**Quasi Monte Carlo.** Quasi Monte Carlo (QMC) methods retain most of the advantages of standard MC while improving the convergence order with respect to the number of samples. At the heart of QMC for estimating expectations is the reinterpretation/approximation of the expected value as an integral with respect to the uniform distribution over the unit hypercube:

$$\mathbb{E}[P] = \int_{\Omega} P(\omega) d\mathbb{P}(\omega) \approx \int_{[0,1]^s} Y(\boldsymbol{x}) d\boldsymbol{x}, \quad (2.5)$$

for some suitable function  $Y$ . QMC methods are, in fact, nothing but quadrature rules over the unit hypercube with  $N$  points and equal weights, approximating the integral on the right-hand side as

$$I = \int_{[0,1]^s} Y(\boldsymbol{x}) d\boldsymbol{x} \approx \frac{1}{N} \sum_{n=1}^N Y(\boldsymbol{x}_n) = I_N, \quad (2.6)$$

where the  $\boldsymbol{x}_n \in \mathbb{R}^s$  are, unlike in the standard MC case, not chosen at random, but chosen carefully and in a deterministic way so as to cover the unit hypercube well. Informally speaking, the slow convergence order of standard MC comes from the fact that true random points, being independent from each other, may fall on the same regions multiple times, while leaving other portions of the domain completely unsearched (see Figure 2.1 left). As a consequence, we have no information about the integrand in these unexplored regions. QMC methods use point sequences that by construction cover the whole space well (see Figure 2.1 right) and therefore achieve better-than-MC convergence rates with respect to  $N$ .

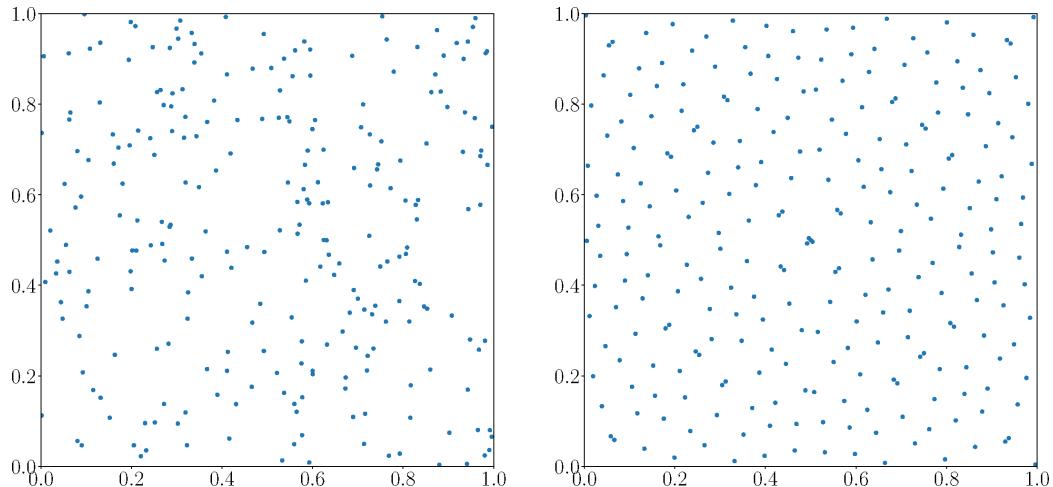


Figure 2.1: *Pseudo-random and QMC point sequences in comparison. On the left, a sample of 256 uniform random points. On the right, the first 256 points in a 2-dimensional Sobol' sequence. It is clear that the points of the Sobol' sequence are spread more uniformly across the unit square.*

**Example 2.1.** According to the probability distribution of  $P$  the approximation sign in equation (2.5) can be replaced by an equality. For instance, if  $P(\omega) = g(Y(\omega))$ , where  $Y(\omega)$  is a standard Gaussian random variable one can write

$$\mathbb{E}[P] = \int_{\mathbb{R}} g(y)f(y)dy = \int_{[0,1]} g(\Phi^{-1}(x)) dx, \quad (2.7)$$

where  $f$  and  $\Phi$  here are the standard normal probability density function (PDF) and cumulative density function (CDF) respectively.

More formally, classical QMC error bounds are in the form (see e.g. [114, 123]),

$$|I - I_N| \leq \mathcal{D}^*(\{\mathbf{x}_n\}_{n=1}^N) \mathcal{V}(Y), \quad (2.8)$$

where  $\mathcal{D}^*(\{\mathbf{x}_n\}_{n=1}^N)$  is the *discrepancy* of the point set  $\{\mathbf{x}_n\}_{n=1}^N$ , which measures how well the point sequence covers the unit hypercube, while  $\mathcal{V}(Y)$  is the *variation* of  $Y$ , and depends on its derivatives.

While random sequences are proven to have discrepancy of  $O((\log \log N/N)^{1/2})$  with probability one [123], hence yielding (up to the logarithmic factor) the standard MC convergence bound, there exist deterministic sequences that are able to achieve discrepancies of  $O((\log N)^s/N)$  [123]. These sequences are called *low-discrepancy* point sequences and, if used for QMC integration, yield a faster-than-MC asymptotic rate of  $O(N^{-1+\epsilon})$ , for any  $\epsilon > 0$ , provided that the integrand  $Y$  is smooth enough. Low-discrepancy sequences include the Sobol' sequences that we use in Chapter 5. Often low-discrepancy sequences must be sampled in powers of 2 (or another basis) or else they lose their good discrepancy properties. For a general description of various QMC point rules we refer to the book by Lemieux [114].

Unlike standard MC, QMC methods are not completely dimension-independent: for high-dimensional problems the  $(\log N)^s$  term in the discrepancy might dominate for small sample sizes. If this happens, low-discrepancy sequences cease to cover the whole hypercube well and their discrepancy temporarily falls back to a  $O(N^{-1/2})$  rate<sup>1</sup> as in the random case up until  $N$  becomes impractically huge [35]. However, this is not always the case. Caflish et al. in [35] investigate this behaviour and introduce the notion of *effective dimensionality*: assume that the QMC integrand  $Y$  can be written as

$$Y = Y^T + Y^C, \quad \text{where } Y \approx Y^T, \quad (2.9)$$

then

$$I = I^T + I^C = \int_{[0,1]^s} Y^T(\mathbf{x})d\mathbf{x} + \int_{[0,1]^s} Y^C(\mathbf{x})d\mathbf{x} \approx I^T. \quad (2.10)$$

---

<sup>1</sup>Caflish et al. also report that QMC integration is almost never worse than standard MC [35].

Further assume that  $Y^T$  either only depends on the first  $s_1 < s$  variables and/or can be expressed as a sum of  $s_2$ -dimensional functions with  $s_2 < s$ , then

$$|I - I_N| \leq |I^T - I_N^T| + |I^C - I_N^C| \leq O\left(\frac{(\log N)^{\bar{s}}}{N}\right) + O\left(\mathcal{V}(Y^C)\frac{(\log N)^s}{N}\right), \quad (2.11)$$

where  $I_N^T$  and  $I_N^C$  are the QMC approximations of  $I^T$  and  $I^C$  respectively and  $\bar{s} = \min(s_1, s_2)$ . If this is the case and  $\bar{s} \ll s$ , then  $Y$  has low effective dimensionality. What this means is that, provided  $\mathcal{V}(Y^C)$  is small, it is possible to replace the  $(\log N)^s$  term with  $(\log N)^{\bar{s}}$ , for which the transition to a  $O(N^{-1})$ -like regime will already happen for small sample sizes. The hope is then that by the time  $N$  becomes large enough so that the second term on the right-hand side in (2.11) is appreciable,  $N$  will also be large enough to observe the asymptotic  $O(N^{-1})$ -like QMC convergence behaviour for this term as well [35].

Not all QMC error bounds are in the form (2.8). In fact, more recent QMC convergence analysis yields error estimates that do not depend on the discrepancy of the point set, but are worst-case error bounds in reproducing kernel Hilbert spaces (RKHS) [47, 109, 110]. These bounds are typically in the form [47]

$$|I - I_N| \leq E(\{\mathbf{x}_n\}_{n=1}^N; \mathcal{H}) \|Y\|_{\mathcal{H}}, \quad E(\{\mathbf{x}_n\}_{n=1}^N; \mathcal{H}) = \sup_{\|f\|_{\mathcal{H}} \leq 1} |I - \frac{1}{N} \sum_{n=1}^N f(\mathbf{x}_n)|, \quad (2.12)$$

where  $\mathcal{H}$  is a suitable RKHS to which  $Y$  belongs. The quantity  $E(\{\mathbf{x}_n\}_{n=1}^N; \mathcal{H})$  is the worst error attainable by the QMC point rule  $\{\mathbf{x}_n\}_{n=1}^N$  over the RKHS  $\mathcal{H}$ . Again, note that  $E$  exclusively depends on the point set and on  $\mathcal{H}$ , but not on  $Y$ , and there exist QMC point rules that achieve faster than  $O(N^{-1/2})$  convergence on suitable RKHSs [47]. The concept of low effective dimensionality also appears in this more recent theory. More specifically, the RKHSs used to achieve convergence are typically weighted, reflecting the underlying assumption that there is “*some varying degree of importance between the variables*” [110]. The great advantage of this worst-case error QMC convergence theory is that it yields error bounds that are dimension-independent (again, in suitable RKHSs).

All things considered, it is clear that for practical applications of high-dimensional QMC integration it is extremely important to order the integration variables in order of decaying importance and/or reduce the dimensionality of the integrand so that higher-than-MC convergence rates can be achieved. This will be a key aspect in the methods we present in Chapter 5.

**Randomised quasi Monte Carlo.** Although theoretically useful, a bound like equation (2.8) cannot be used in practice as the quantities involved (discrepancy and variation measures) are extremely difficult to estimate. The quantities in 2.12 are relatively easier to work with, but their computation is still non-trivial. All in all, QMC point sequences are deterministic and we cannot rely on the central limit theorem as for standard MC.

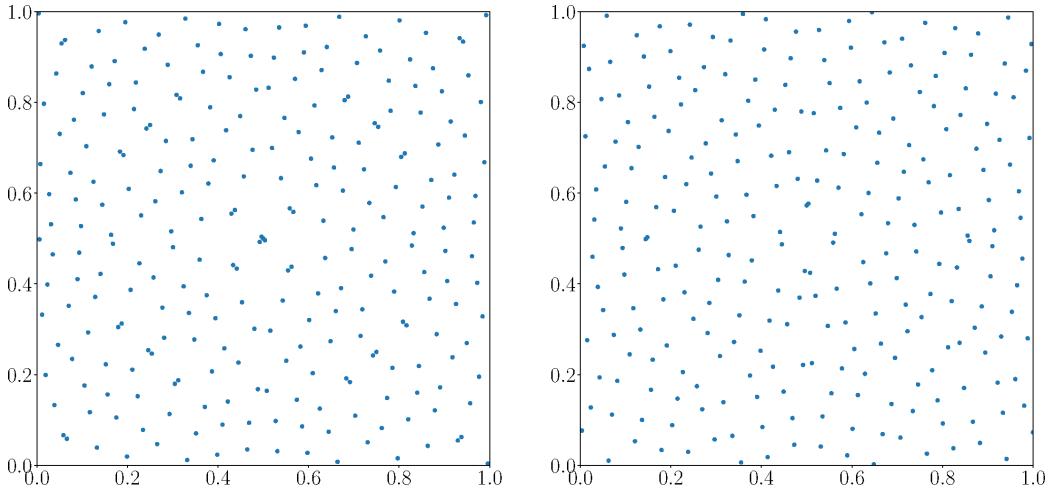


Figure 2.2: *Randomised and not randomised low-discrepancy sequences in comparison.* On the left, the first 256 points in a 2-dimensional Sobol' sequence. On the right, the same points after random digital shifting. The randomised sequence still covers the unit square well.

Randomised QMC methods combine MC and QMC ideas and fix these issues by randomising the QMC point sequence used, i.e. given a fixed deterministic QMC integration rule  $\{\mathbf{x}_n\}_{n=1}^N$ , randomised QMC produces a set of  $M$  independent randomised sequences  $\{\hat{\mathbf{x}}_{n,m}(\omega)\}_{n=1,m=1}^{N,M}$  so that either 1) the discrepancy of the point set is preserved (see Figure 2.2) [114], or 2) mean-square error bounds in the RKHS with respect to the randomised sequence can be obtained [47]. The main randomisation techniques are random shifting and scrambling; see Chapter 6 of [114] for an overview and [137] for a comparison of various techniques. The randomised sequences are then combined into the randomised QMC estimator,

$$\hat{I}_{M,N}(\omega) = \frac{1}{M} \sum_{m=1}^M I_N^m(\omega) = \frac{1}{M} \sum_{m=1}^M \left( \frac{1}{N} \sum_{n=1}^N Y(\hat{\mathbf{x}}_{n,m}(\omega)) \right). \quad (2.13)$$

Since  $I_N^m(\omega)$  is now random, provided that  $M$  is large enough a confidence interval can be estimated and we retain a practical error measure as in the standard MC case. Note that, however, while the total cost of randomised QMC is proportional to  $NM$ , the asymptotic convergence order of the randomised QMC estimator is  $O(\log(N)^s/(NM^{1/2}))$  and therefore it is always more advantageous to increase  $N$  rather than  $M$ , and  $M$  should be kept as small as possible. Typical choices for  $M$  range between 8 and 128. However, extra care must be taken with small values of  $M$  when computing confidence intervals as in this case  $\hat{I}_{M,N}$  is not approximately normal and the central limit theorem cannot be applied. In this work we use  $M = 32$  unless otherwise stated. Assuming fixed  $M$ , a given  $MSE$  tolerance  $\varepsilon^2$ , a  $O(\varepsilon^{-q})$  cost per sample and a QMC convergence order of  $O(N^{-1+\epsilon})$  for any  $\epsilon > 0$ , the total cost of randomised QMC is  $O(\varepsilon^{-q-1/(1-\epsilon)})$ , which for small  $\epsilon$  is almost  $\varepsilon^{-1}$  times better than standard MC.

*Remark 2.1.* Let us quickly compare standard MC and randomised QMC: 1) QMC convergence order and total cost are superior than MC; 2) standard MC has almost

no smoothness requirements, while QMC requires the mixed derivative of  $Y$  to be in  $L^1([0, 1]^s)$ ; 3) neither MC nor suitable QMC rules suffer from the curse of dimensionality when performing high-dimensional integration; 4) each MC and QMC sample can be parallelised; 5) both MC and randomised QMC provide confidence intervals for the solution.

In the rest of the thesis we will refer to randomised QMC as just QMC.

## 2.2 Multilevel Monte Carlo methods

While using QMC point sequences reduces the computational cost of MC by increasing the convergence speed, multilevel Monte Carlo methods can make the cost of computing one sample of  $P$   $O(1)$  on average, i.e. multilevel methods are, under suitable conditions, able to get rid of the  $O(\varepsilon^{-q})$  term in the MC cost complexity. From applying a multilevel strategy to standard Monte Carlo we obtain the multilevel Monte Carlo method (MLMC) [69], which can achieve  $O(\varepsilon^{-2})$  complexity. In certain cases, it is possible to get the best of two worlds by combining QMC with MLMC, yielding the multilevel quasi Monte Carlo method (MLQMC) [67] with complexity  $O(\varepsilon^{-1/\chi})$ , where  $\chi \geq 1/2$ .

### 2.2.1 Multilevel Monte Carlo

The multilevel Monte Carlo (MLMC) was first introduced by Heinrich in [87] for parametric integration and popularised by Giles for stochastic differential equations (SDEs) in [69]. Assume that it is possible to compute realisations of  $P(\omega)$  at different accuracy levels  $P_\ell(\omega)$  for  $\ell = 1, \dots, L$  of increasing accuracy and computational cost and that the approximation of  $P$  on the finest level,  $P_L$  is accurate enough. The MLMC method estimates  $\mathbb{E}[P_L]$  through the telescopic sum,

$$\mathbb{E}[P] \approx \mathbb{E}[P_L] = \sum_{\ell=1}^L \mathbb{E}[P_\ell - P_{\ell-1}], \quad (2.14)$$

where  $P_0 \equiv 0$ . Each of the terms in the sum can be separately approximated by using standard Monte Carlo. This yields the MLMC estimator:

$$\mathbb{E}[P_L] \approx \sum_{\ell=0}^L \left[ \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} (P_\ell(\omega_\ell^n) - P_{\ell-1}(\omega_\ell^n)) \right]. \quad (2.15)$$

In a finite element context, the levels of accuracy can be defined by using a hierarchy of meshes ( $h$ -refinement) or by increasing the polynomial degree of the finite elements used ( $p$ -refinement).

Let  $C_\ell$ ,  $V_\ell$  be the cost and variance of one sample  $P_\ell(\cdot) - P_{\ell-1}(\cdot)$  respectively. Then, the total MLMC cost and variance are

$$C_{tot} = \sum_{\ell=1}^L N_\ell C_\ell, \quad \hat{V} = \sum_{\ell=0}^L N_\ell^{-1} V_\ell. \quad (2.16)$$

We can minimise the estimator variance for fixed total cost. For further details we refer to [68]. This gives that, for a fixed MSE tolerance  $\varepsilon^2$ , the optimal number of samples for each level and related total cost are,

$$N_\ell = \left( \varepsilon^{-2} \sum_{l=1}^L \sqrt{V_l C_l} \right) \sqrt{V_\ell / C_\ell}, \quad C_{tot} = \varepsilon^{-2} \left( \sum_{\ell=1}^L \sqrt{V_\ell C_\ell} \right)^2. \quad (2.17)$$

*Remark 2.2.* We now compare the cost complexity of standard and multilevel Monte Carlo for the estimation of  $\mathbb{E}[P_L]$ . Let  $\mathbb{V}[P_L] = O(V_1)$ , then the total cost complexity of standard MC is  $O(\varepsilon^{-2} V_1 C_L)$ . According to how the product  $V_\ell C_\ell$  varies with level, we can have three different scenarios for MLMC:

1. The product  $V_\ell C_\ell$  increases with level. Then, to leading order, the total MLMC cost is  $O(\varepsilon^{-2} V_L C_L)$ , for an improvement in computational cost over standard Monte Carlo by a  $V_1/V_L$  factor.
2. The product is constant with the level. Then, we have a MLMC total cost of  $O(\varepsilon^{-2} L^2 V_L C_L) = O(\varepsilon^{-2} L^2 V_1 C_1)$ . This gives a cost improvement of  $V_1/(L^2 V_L) \approx C_L/(L^2 C_1)$  with respect to standard MC.
3. The product decreases with the level. Then,  $C_{tot} \approx O(\varepsilon^{-2} V_L C_1)$ , for an improvement of  $C_L/C_1$ . For example this could be the ratio between a fine mesh PDE solution cost and a coarse mesh PDE solution cost, which is generally quite significant.

*Remark 2.3.* Note that even with MLMC we cannot improve over the  $N = O(\varepsilon^{-2})$  samples required for a given accuracy, so we still need to run many PDE solves. In case we have a good hierarchy of levels available, however, we can significantly reduce the computational cost.

The MSE of the MLMC estimator is given by,

$$MSE = \hat{V} + E[\hat{P} - P]^2, \quad (2.18)$$

where  $\hat{P}$  is the MLMC estimator of variance  $\hat{V}$ . To ensure that  $MSE \leq \varepsilon^2$ , we enforce the bounds,

$$\hat{V} \leq (1 - \theta)\varepsilon^2, \quad E[\hat{P} - P]^2 \leq \theta\varepsilon^2, \quad (2.19)$$

where  $\theta \in (0, 1)$  is a weight, introduced by Haji-Ali et al. in [82], that balances the two terms so as to obtain comparable error reduction.

The convergence of MLMC is ensured by the following theorem.

**Theorem 2.1** ([68], theorem 1). *Let  $P \in L^2(\Omega, \mathbb{R})$  and let  $P_\ell$  be its level  $\ell$  approximation. Let  $Y_\ell$  be the MC estimator of  $\mathbb{E}[P_\ell - P_{\ell-1}]$  on level  $\ell$  such that*

$$\mathbb{E}[Y_\ell] = \mathbb{E}[P_\ell - P_{\ell-1}], \quad (2.20)$$

with  $P_0 = 0$ , and let  $C_\ell$  and  $V_\ell$  be the expected cost and variance of each of the  $N_\ell$  Monte Carlo samples needed to compute  $Y_\ell$ . If the estimators  $Y_\ell$  are independent and there exist positive constants  $\alpha, \beta, \gamma, c_1, c_2, c_3$ , such that  $\alpha \geq \frac{1}{2} \min(\beta, \gamma)$  and

$$|\mathbb{E}[P_\ell - P]| \leq c_1 2^{-\alpha\ell}, \quad V_\ell \leq c_2 2^{-\beta\ell}, \quad C_\ell \leq c_3 2^{\gamma\ell}, \quad (2.21)$$

then there exists a positive constant  $c_4$  such that, for all  $\varepsilon < e^{-1}$ , there is a level number  $L$  and number of samples  $N_\ell$ , such that the MLMC estimator

$$\hat{P} = \sum_{\ell=1}^L Y_\ell, \quad (2.22)$$

has MSE with bound,

$$MSE = \mathbb{E}[(\hat{P} - \mathbb{E}[P])^2] \leq \varepsilon^2, \quad (2.23)$$

with a total computational complexity with bound,

$$\mathbb{E}[C_{tot}] \leq \begin{cases} c_4 \varepsilon^{-2}, & \beta > \gamma, \\ c_4 \varepsilon^{-2} (\log \varepsilon)^2, & \beta = \gamma, \\ c_4 \varepsilon^{-2-(\gamma-\beta)/\alpha}, & \beta < \gamma. \end{cases} \quad (2.24)$$

*Remark 2.4.* Usually, in practice, the values of  $\alpha, \beta, \gamma, c_i$  for  $i = 1, 2, 3$ , are not known and they must be estimated. In the case in which  $P$  is an output functional of the solution of a PDE with random coefficients solved with the finite element method tentative values can be derived using *a priori/a posteriori* error estimates [31] and iterative solver complexity results [46]. However, it is also possible to just estimate  $V_\ell$  and  $C_\ell$ , and use relation (2.17) to determine the optimal number of samples needed [68].

We can finally present the MLMC algorithm:

### MLMC algorithm (taken from [68])

- Set the required tolerance  $\varepsilon$ , the maximum level  $L_{\max}$ , the initial number of levels  $L$  and the initial number of samples  $\bar{N}_\ell$  to be taken on each of them.
- **while** extra samples need to be evaluated ( $\exists \ell : \bar{N}_\ell > 0$ ):
  1. for each level, evaluate all the samples scheduled to be taken;
  2. compute/update estimates for the level variance  $V_\ell$ ,  $\ell = 1, \dots, L$ ;
  3. compute optimal number of samples  $N_\ell$  by using (2.17) and update the numbers of extra samples needed  $\bar{N}_\ell$  accordingly;
  4. test for weak convergence, i.e. check whether the weak error  $\mathbb{E}[\hat{P} - P]^2$  is below the required tolerance;
  5. if not converged: if  $L = L_{\max}$  report failed convergence; otherwise set  $L := L + 1$ , update  $N_\ell$  and  $\bar{N}_\ell$  and compute  $N_L = \bar{N}_L$ .

### 2.2.2 Multilevel quasi Monte Carlo

The idea of multilevel quasi Monte Carlo, originally presented in [67], is to combine QMC and MLMC together with the objective of also combining their advantages. Again, let  $\{P_\ell\}_{\ell=1}^L$  be a multilevel hierarchy of approximations. MLQMC also estimates  $\mathbb{E}[P]$  through the same telescopic sum as MLMC,

$$\mathbb{E}[P] \approx \mathbb{E}[P_L] = \sum_{\ell=1}^L \mathbb{E}[P_\ell - P_{\ell-1}], \quad P_0 \equiv 0. \quad (2.25)$$

The difference between MLMC and MLQMC is then in the way each expectation in the sum is approximated: MLMC uses standard MC, MLQMC uses randomised QMC. The MLQMC estimator is in fact obtained by approximating each term with

$$\mathbb{E}[P_\ell - P_{\ell-1}] = \int_{[0,1]^{s_\ell}} Y_\ell(\mathbf{x}) d\mathbf{x} \approx \frac{1}{M} \sum_{m=1}^M \left( \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} Y_\ell(\hat{\mathbf{x}}_{n,m}^\ell(\omega)) \right) = \frac{1}{M} \sum_{m=1}^M I_{N_\ell}^{m,\ell}(\omega), \quad (2.26)$$

where the meaning of each variable is the same as in the QMC case. The difference compared with QMC is that now we have a hierarchy of integrands  $\{Y_\ell\}_{\ell=1}^L$  and of randomised QMC point sequences  $\{\hat{\mathbf{x}}_{n,m}^\ell\}_{n=1,m=1,\ell=1}^{N_\ell,M,\ell=L}$  of dimensions  $\{s_\ell\}_{\ell=1}^L$ .

In principle, the convergence theory would carry through in roughly the same way as in MLMC, i.e. by finding the optimal level and number of samples to achieve a given MSE error tolerance. However, in the MLQMC case there is a key complication that makes the analysis extremely complicated: the  $Y_\ell$  for  $\ell > 1$  represent the difference between two different approximations and as such, might not present the same smoothness as  $Y$  in the QMC only case. In fact, Giles and Waterhouse report in [67] that for some financial SDE applications the QMC convergence rate decays to  $O(N^{-1/2})$  and/or transitions later to a higher QMC rate as  $\ell$  grows. What could be happening is that the difference between the approximations on different levels becomes rougher and rougher making  $Y_\ell$  either non-smooth or of high effective dimensionality. The same behaviour does not necessarily occur in PDE applications [109].

In general, proving any convergence result for MLQMC is particularly hard, to the extent that convergence proofs are only available for a few specific problems and specific QMC point sequences [88, 109]. For this reason, setting up a MLQMC estimator optimally can be challenging and it might not even be clear a priori whether resorting to a pure MLMC or QMC strategy is actually more advantageous (in the QMC case we only deal with  $Y$  which is generally smoother) and the various methods have to be tested. However, in the best possible case where we get a  $O(N^{-\chi})$ ,  $1/2 \leq \chi \leq 1$ , QMC rate for each term in the telescoping sum, the benefits of MLMC and QMC can accumulate yielding a total MLQMC computational cost of  $O(\varepsilon^{-1/\chi})$  for a given MSE tolerance of  $\varepsilon^2$  [88]. In this case MLQMC significantly outperforms all the other Monte Carlo methods.

We conclude this section by presenting the MLQMC algorithm we will be using in this thesis: let  $C_\ell$  be the cost of evaluating  $Y_\ell$  and let  $V_\ell = \mathbb{V}[I_{N_\ell}^{m,\ell}]$ ,

### MLQMC algorithm (taken from [67])

1. Set the required tolerance  $\varepsilon, \theta \in (0, 1)$ , the minimum and maximum level  $L_{\min}$  and  $L_{\max}$  and the initial number of levels to be  $L = 1$ .
2. Get an initial estimate of  $V_L$  with  $N_L = 1$  and  $M = 32$  randomisations.
3. While  $\sum_{\ell=1}^L V_\ell > (1 - \theta)\varepsilon^2$ , double  $N_\ell$  on the level with largest  $V_\ell/(C_\ell N_\ell)$ .
4. If  $L < L_{\min}$  or the bias estimate is greater than  $\sqrt{\theta}\varepsilon$ , set  $L = L + 1$ . If  $L \leq L_{\max}$  go to 2, otherwise report convergence failure.

*Remark 2.5* (Adapted from [67]). The term  $\sum_{\ell=1}^L V_\ell$  is the total estimator variance and the variable  $\theta$  is a weight with the same meaning as in the MLMC case. The choice of  $N_\ell$  on each level is heuristic: doubling the number of samples will eliminate (independently on whether we are in an MC or QMC-like convergence rate regime) most of the estimated variance  $V_\ell$  on level  $\ell$  at a cost  $N_\ell C_\ell$  and we therefore double the number of samples on the level that offers the largest variance reduction per unit cost.

## 2.3 Standard and generalised stochastic fields

In this section we highlight the few notions of functional and stochastic analysis needed in the rest of the thesis. First of all, we adopt the following notation and definitions.

**Subsets of compact closure.** Given an open domain  $G \subseteq D$ , we write  $G \subset\subset D$  to indicate that the closure of  $G$  is a compact subset of  $D$ .

**$L^2$  inner product.** For an open domain  $D \subseteq \mathbb{R}^d$ , we let  $(\cdot, \cdot)$  denote the  $L^2(D)$  inner product where  $L^2(D)$  is the standard Hilbert space of square-integrable functions on  $D$ .

**Sobolev spaces.** For an open domain  $D \subseteq \mathbb{R}^d$ , we indicate with  $W^{s,q}(D)$ ,  $s > 0$ ,  $q \geq 1$  the Sobolev space of integrability order  $q$  and differentiability order  $s$ , and we indicate with  $W_{\text{loc}}^{s,q}(D)$  the space of functions which are  $W^{s,q}(G)$ , for all  $G \subset\subset D$ . Furthermore, we let  $H^s(D) \equiv W^{s,2}(D)$  and  $H_{\text{loc}}^s(D) \equiv W_{\text{loc}}^{s,2}(D)$  and we denote with  $H^{-s}(D)$  the dual space of  $H_0^s(D)$ , where  $H_0^s(D)$  is the space of  $H^s(D)$  functions with zero trace on  $\partial D$ .

**Hölder spaces.** For an open domain  $D \subseteq \mathbb{R}^d$ , we indicate with  $C^k(D)$ , where  $k \in \mathbb{N}$  the space of  $k$ -times continuously differentiable functions defined on  $D$  and with  $C^{k,\alpha}(D)$ , with  $\alpha \in (0, 1]$  the space of  $C^k(D)$  functions that have  $\alpha$ -Hölder continuous  $k$ -th partial derivatives.

In this thesis we will be working with PDEs with random coefficients, where the coefficients will not be mere random variables, but stochastic (or random) fields, defined as follows.

**Definition 2.1** (Stochastic field [3]). Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space. A stochastic (or random) field  $X = X(\mathbf{x}, \omega)$ ,  $\mathbf{x} \in \mathbb{R}^d$ ,  $\omega \in \Omega$ , is a function whose values are

random variables for each  $\mathbf{x} \in \mathbb{R}^d$ . Given a Banach space  $U$  we say that a field  $X \in L^2(\Omega, U)$  if  $X(\mathbf{x}, \cdot) \in L^2(\Omega, \mathbb{R})$  for almost every  $\mathbf{x} \in D$  and  $X(\cdot, \omega) \in U$  almost surely (a.s.).

A stochastic field is a function of space for fixed  $\omega$  and the values of a stochastic field are square integrable random variables in  $L^2(\Omega, \mathbb{R})$ . If these random variables are all jointly Gaussians, we then have a Gaussian random field:

**Definition 2.2** (Gaussian random field [3]). A stochastic field is Gaussian if the random variables that define its values are all joint Gaussian. A Gaussian field is uniquely determined by providing a mean function  $\mu(\mathbf{x})$  and a symmetric positive definite covariance function  $C(\mathbf{x}, \mathbf{y})$ . A Gaussian field is stationary if  $\mu$  is constant and if the covariance only depends on the difference  $\mathbf{x} - \mathbf{y}$ , i.e. if  $C = C(\mathbf{x} - \mathbf{y})$ . A Gaussian field is isotropic if it is stationary and the covariance only depends on  $r = \|\mathbf{x} - \mathbf{y}\|_2$ , i.e. if  $C = C(r)$ .

A very important family of Gaussian fields that we will use extensively in our work is the Matérn class,

**Definition 2.3** (Matérn random field [3]). A Matérn field is a Gaussian field with covariance of the Matérn class, i.e. in the form

$$C(\mathbf{x}, \mathbf{y}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa r)^\nu \mathcal{K}_\nu(\kappa r), \quad r = \|\mathbf{x} - \mathbf{y}\|_2, \quad \kappa = \frac{\sqrt{8\nu}}{\lambda}, \quad \mathbf{x}, \mathbf{y} \in D, \quad (2.27)$$

where  $\sigma^2, \nu, \lambda > 0$  are the variance, smoothness parameter and correlation length of the field respectively,  $\Gamma(x)$  is the Euler Gamma function and  $\mathcal{K}_\nu$  is the modified Bessel function of the second kind.

The correlation length  $\lambda$  roughly represents the distance beyond which point values of the field are approximately uncorrelated while the parameter  $\nu$  is called the smoothness parameter as it is directly related to the smoothness of the field, as stated in the following result,

**Theorem 2.2** (Smoothness of a Matérn field [3, 150]). Let  $u(\mathbf{x}, \omega)$  be a zero-mean Gaussian field with covariance function given by (2.27), then,

$$u(\cdot, \omega) \in H_{loc}^{\nu-\epsilon}(\mathbb{R}^d) \cap C^{\lceil \nu-1 \rceil}(\mathbb{R}^d), \quad \forall \epsilon > 0, \quad a.s. \quad (2.28)$$

The above definitions are well-known. We now introduce the concept of *generalised stochastic fields* in a less standard way: as it turns out there are two non-equivalent definitions that can be found in the literature.

**Definition 2.4** (Generalised stochastic field (type I)). Let  $V$  be a Banach space. Following the definition introduced by Itô [105] and extended by Inaba and Tapley [95] we denote with  $\mathcal{L}(V, L^2(\Omega, \mathbb{R}))$  the space of generalised stochastic fields (of type I) that are continuous linear mappings from  $V$  to  $L^2(\Omega, \mathbb{R})$ . For a given  $\xi \in \mathcal{L}(V, L^2(\Omega, \mathbb{R}))$  we indicate the action (or pairing) of  $\xi$  on a function  $\phi \in V$  with the notation  $\xi(\phi) = \langle \xi, \phi \rangle$ .  $\mathcal{L}(V, L^2(\Omega, \mathbb{R}))$  is a Banach space with norm

$$\|\xi\|_{\mathcal{L}(V, L^2(\Omega, \mathbb{R}))} = \sup_{v \in V, \|v\|_V \leq 1} \mathbb{E}[|\langle \xi, v \rangle|^2]^{1/2}. \quad (2.29)$$

*Remark 2.6* (Angle bracket notation). In this thesis we denote by  $\langle \xi, \phi \rangle$  the action of a type I stochastic field against a test function  $\phi$  since we believe that this notation is more convenient than  $\dot{W}(\phi)$ . The angle bracket notation is often used in the literature for duality pairings, but this is not the case in this thesis.

Another definition of generalised stochastic fields which has been adopted in the literature [140] is the following.

**Definition 2.5** (Generalised stochastic field (type II)). Let  $U$  be a Banach space with dual  $U^*$ . A mapping  $\xi : U \times \Omega \rightarrow \mathbb{R}$  is a generalised stochastic field (of type II) if for all  $\phi \in U$ ,  $\xi(\phi, \cdot)$  is a real valued random variable and for all  $\omega \in \Omega$ ,  $\xi(\cdot, \omega)$  is an element in  $U^*$ . We denote with  $L^2(\Omega, U^*)$  the space of generalised stochastic fields of type II that are elements of  $U^*$  almost surely such that for all  $\phi \in U$ ,  $\xi(\phi, \cdot) \in L^2(\Omega, \mathbb{R})$ .  $L^2(\Omega, U^*)$  is a Banach space with norm

$$\|\xi\|_{L^2(\Omega, U^*)} = \mathbb{E} \left[ \left( \sup_{v \in U, \|v\|_U \leq 1} |\xi(v, \cdot)| \right)^2 \right]^{1/2} = \mathbb{E}[\|\xi\|_{U^*}^2]^{1/2}. \quad (2.30)$$

*Remark 2.7* (Non-equivalence of the two definitions of generalised random fields). The two definitions of generalised stochastic fields are not equivalent, i.e. a generalised stochastic field of type I is not necessarily a generalised stochastic field of type II as well and vice versa [140]. Furthermore, even if a generalised stochastic field is of both types so that  $\xi \in \mathcal{L}(V, L^2(\Omega, \mathbb{R}))$  and  $\xi \in L^2(\Omega, U^*)$  the spaces  $V$  and  $U$  do not need to be the same (see the work by Pilipović and Seleši [140]). An example of such a field is spatial white noise (defined next). This issue does not arise for proper stochastic fields: all stochastic fields are also generalised stochastic fields of both types and  $V \equiv U$  [140].

Possibly the most commonly used generalised random field is Gaussian *white noise*. White noise is going to be extremely important throughout this thesis and is defined as follows.

**Definition 2.6** (White noise, see example 1.2 and lemma 1.10 in [89]). Let  $D \subseteq \mathbb{R}^d$  be an open domain. The white noise  $\dot{W} \in \mathcal{L}(L^2(D), L^2(\Omega, \mathbb{R}))$  is a generalised stochastic field of type I such that for any collection of  $L^2(D)$  functions  $\{\phi_i\}$ , if we let  $b_i = \langle \dot{W}, \phi_i \rangle$ , then  $\{b_i\}$  are joint Gaussian random variables with zero mean and covariance given by  $\mathbb{E}[b_i b_j] = (\phi_i, \phi_j)$ .

*Remark 2.8.* White noise is also a generalised stochastic field of type II (see [182]). More specifically, for any  $\epsilon > 0$  we have

$$\dot{W} \in \mathcal{L}(L^2(D), L^2(\Omega, \mathbb{R})) \quad \text{and} \quad \dot{W} \in L^2(\Omega, H^{-d/2-\epsilon}(D)). \quad (2.31)$$

Drawing samples of (generalised) random fields from their distribution is in general a computationally challenging task. We conclude this background chapter with an overview of computational methods for Gaussian field sampling.

## 2.4 Gaussian field sampling

Some of the contents of this section are adapted from our paper [43].

In this thesis, we mainly work with Gaussian fields and white noise, where the Gaussian fields appear as coefficients of PDEs to be solved with Monte Carlo methods. This requires the efficient sampling of Gaussian fields from their distribution. Unfortunately, while sampling independent Gaussian random variables is “easy” as fast pseudo-random number generators are available in many software packages, sampling a Gaussian field with a prescribed correlation function is much harder.

In what follows we give an overview of existing Gaussian field sampling methods and we then describe in further detail the SPDE approach, which we use throughout the thesis. We give an overview of the existing sampling methods in Table 2.1.

Method	Set-up	Memory	Sample	When does it work well?
direct factorisation	$m^3$	$m^2$	$m^2$	small problem sizes
pivoted Cholesky	$m_{\text{PC}}^2 m$	$m_{\text{PC}} m$	$m_{\text{PC}} m$	small problem sizes or smooth $C(\mathbf{x}, \mathbf{y})$
H-matrix approx.	$m \log m$	$m \log m$	$m \log m$	small problem sizes or smooth $C(\mathbf{x}, \mathbf{y})$ (†)
$\text{H}^2$ -matrix approx.	$m$	$m$	$m \log m$	small problem sizes or smooth $C(\mathbf{x}, \mathbf{y})$ (†)
KL expansion	$m_{\text{KL}}^2 m$	$m_{\text{KL}} m$	$m_{\text{KL}} m$	small problem sizes or smooth $C(\mathbf{x}, \mathbf{y})$
circulant embedding	$s \log s$	$s$	$s \log s$	Stationary fields on structured grids.
SPDE approach	$s$	$s$	$ks (*)$	Matérn fields with small $\nu = 2k - d/2$ .

Table 2.1: A complexity comparison between the main Gaussian field sampling methods available in the literature. All complexity estimates in the table are to be understood in the  $O(\cdot)$  sense. We refer to the rest of this section for the presentation of the various methods. KL stands for Karhunen-Loëve. Both the circulant embedding and the SPDE approach require sampling the random field on an extended domain where  $s = s(m)$  point values of the field are needed. The  $k$  in the H-matrix approximation row is the rank of the H-matrix approximation, while the  $k$  in the SPDE approach row is the  $k$  such that  $\nu = 2k - d/2$ . (\*) If  $k$  is a positive integer, then the sample cost is only  $O(ks)$ , if  $k$  is non-integer, then an extra  $O(s \log s)$  term is needed and the sample complexity is  $O(ks + s \log s)$  [27]. (†) For the H/ $\text{H}^2$ -matrix approach it is sufficient for  $C(\mathbf{x}, \mathbf{y})$  to be asymptotically smooth [62].

### 2.4.1 Matrix factorisations

Let  $u$  be the Gaussian field we are interested in sampling, which we assume to have zero mean without loss of generality. In practice, samples of  $u$  are always only needed at discrete locations  $\mathbf{x}_1, \dots, \mathbf{x}_m \in D$ . The perhaps simplest sampling strategy consists in drawing realisations of a Gaussian vector  $\mathbf{u} \sim \mathcal{N}(0, C)$  with  $\mathbf{u}_i = u(\mathbf{x}_i)$  and covariance matrix  $C_{ij} = \mathbb{E}[u(\mathbf{x}_i)u(\mathbf{x}_j)]$ . This approach is usually computationally expensive as it requires the factorisation of the dense covariance matrix  $C \in \mathbb{R}^{m \times m}$ . In fact, if we let  $\mathbf{z} \sim \mathcal{N}(0, I)$  be a standard Gaussian vector and we factorise  $C = HH^T$

with  $H \in \mathbb{R}^{m \times n}$ , we can sample  $\mathbf{u}$  as  $\mathbf{u} = H\mathbf{z}$  since,

$$\mathbb{E}[\mathbf{u}\mathbf{u}^T] = \mathbb{E}[H\mathbf{z}(H\mathbf{z})^T] = H\mathbb{E}[\mathbf{z}\mathbf{z}^T]H^T = HIH^T = C. \quad (2.32)$$

A basic form of this method uses the Cholesky factorisation of  $C$ . In this case  $H$  is square and dense lower-triangular and the factorisation has a computational complexity of  $O(m^3)$ , yielding an  $O(m^3)$  setup cost and an  $O(m^2)$  storage requirement and cost per sample.

*Remark 2.9.* Usually  $n$  is taken to be equal to  $m$  so that  $H \in \mathbb{R}^{m \times n}$  is square. However this is not necessary for (2.32) to hold. For instance, one of the sampling strategies we present in this thesis uses  $n > m$ .

If the field is smooth enough so that the eigenvalues of  $C$  are rapidly decaying this method can be made competitive by using a low-rank approximation of rank  $m_{\text{PC}}$  instead (this can be interpreted as a discrete version of the Karhunen-Loëve expansion which we present next). Common choices are either a truncated singular value decomposition (SVD) or the pivoted Cholesky factorisation [84]. For a fixed tolerance  $\varepsilon$ , the pivoted Cholesky decomposition computes a low rank approximation  $HH^T \approx C$  of  $C$ , where  $H \in \mathbb{R}^{m \times m_{\text{PC}}}$ , such that  $\text{trace}(C - HH^T) \leq \varepsilon$ . This gives an overall factorisation cost of  $O(m_{\text{PC}}^2 m)$  and cost per sample and storage requirement of  $O(m_{\text{PC}} m)$ , where  $m_{\text{PC}} \leq m$ . If the covariance function  $C(\mathbf{x}, \mathbf{y})$  is smooth and with large enough correlation length (e.g. Matérn covariance with  $\nu = \infty$  and  $\lambda$  large), then the eigenvalues of  $C$  decay quickly and a truncated decomposition with  $m_{\text{PC}} \ll m$  can be obtained, yielding a fast sampling method. If, however, the covariance is not smooth and/or the correlation length is small (e.g. Matérn covariance with  $\nu = 1/2$  and/or small  $\lambda$ ), the covariance matrix is effectively high rank and the  $m_{\text{PC}}$  needed by the algorithm will effectively be very close to  $m$  unless the application at hand can work with a relatively high error tolerance  $\varepsilon$ .

A generalisation of the pivoted Cholesky approach consists of approximating the covariance matrix with the closest rank  $k$  hierarchical matrix (H-matrix or H<sup>2</sup>-matrix), which can then be factorised efficiently. This technique consists of hierarchically dividing  $C$  into sub-blocks that are then approximated by low-rank matrices (see [62, 102, 119] for some applications to Gaussian field sampling). Note that approximation of  $C$  by H/H<sup>2</sup>-matrices might destroy its symmetry and positive-definiteness and extra care must be taken [81].

The hierarchical matrix approach works well under the assumption that the covariance function is asymptotically smooth, i.e.  $C^\infty(D \times D \setminus \{\mathbf{x} = \mathbf{y}\})$  with additional conditions on the derivatives at  $x = y$  and at infinity (see [50, 62]; Matérn fields satisfy this requirement). In the H-matrix case, this assumption leads to an overall offline and per-sample cost and storage complexity of  $O(m \log m)$  [50, 81]. In the H<sup>2</sup>-matrix case, two strategies are available: it is either possible to factorise the H<sup>2</sup>-matrix approximation or to not factorise it and exploit the fast matrix-vector products given by the matrix format within an iterative solver to compute the action of the matrix square root instead. In the former case, this yields a  $O(m \log m)$  offline cost and a  $O(m)$  memory and cost per-sample complexity. However, as reported by Feischl et al. in [62], there is no complete error analysis available for this approach yet. In the latter

case, we have a  $O(m)$  offline cost and memory cost and a  $O(m \log m)$  per-sample cost complexity, but precise theoretical error estimates are available [62] (hence why these are the complexity estimates reported in Table 2.1). The H/H<sup>2</sup>-matrix approximation error analysis is quite involved and results are only available for a few particular cases [50, 62, 81]. We remark that the accuracy and performance of this approach deteriorates if the covariance lacks the required smoothness [81, 102].

Overall, the main advantage of these factorisation approaches is the flexibility: they work for any collection of evaluation points  $\mathbf{x}_1, \dots, \mathbf{x}_m$  and for any covariance function. The downside is that the computational cost can be prohibitive when  $C(\mathbf{x}, \mathbf{y})$  is non-smooth.

### 2.4.2 Karhunen-Loëve expansion

Instead of computing a factorisation of the covariance matrix, it is also possible to work at the continuous level and expand the field  $u$  as a (possibly finite or truncated) series of basis functions. Different choices of bases yield different methods. Common choices are the basis of the eigenfunctions of  $C(\mathbf{x}, \mathbf{y})$  (Karhunen-Loëve expansion), the Fourier basis (circulant embedding [49]), or a finite element basis ([27, 51, 116, 182]).

We first discuss the Karhunen-Loëve (KL-) expansion approach (adapted from Sullivan's book [163]). Let  $\mathcal{C}$  be the covariance operator of  $u$ , defined as follows: for any  $v \in L^2(D)$ ,

$$(\mathcal{C}v)(\mathbf{x}) := \int_D C(\mathbf{x}, \mathbf{y})v(\mathbf{y})d\mathbf{y}, \quad (2.33)$$

where  $C(\mathbf{x}, \mathbf{y})$  is the covariance function of  $u$ . The idea of the Karhunen-Loëve approach is to expand  $u$  as a series of eigenfunctions of  $\mathcal{C}$ . Before we introduce this expansion, it is useful to state Mercer's theorem as it describes the properties of the eigenvalue-eigenfunction pairs of  $\mathcal{C}$ .

**Theorem 2.3** (Mercer, adapted from [163]). *Let  $D \subset\subset \mathbb{R}^d$  and let  $C(\mathbf{x}, \mathbf{y}) : D \times D \rightarrow \mathbb{R}$  be a continuous symmetric positive semi-definite covariance function with associated covariance operator  $\mathcal{C}$ . Then there exists a family  $\{(\lambda_i, \phi_i)\}_{i \in \mathbb{N}}$  of eigenpairs of  $\mathcal{C}$  such that*

$$\int_D C(\mathbf{x}, \mathbf{y})\phi_i(\mathbf{y})d\mathbf{y} = \lambda_i\phi_i(\mathbf{x}), \quad (\phi_i, \phi_j) = \delta_{ij}, \quad (2.34)$$

( $\delta_{ij}$  here is the Kronecker delta) where the  $\lambda_i$  are non-negative, the  $\phi_i$  form an orthonormal basis of  $L^2(D)$ , and the eigenfunctions corresponding to non-zero eigenvalues are continuous. Furthermore,

$$C(\mathbf{x}, \mathbf{y}) = \sum_{n \in \mathbb{N}} \lambda_n \phi_n(\mathbf{x}) \phi_n(\mathbf{y}), \quad (2.35)$$

and the series converges absolutely and uniformly over compact subsets of  $D$ .

We can now state Karhunen-Loëve's theorem:

**Theorem 2.4** (Karhunen-Loëve, adapted from [163]). *Let  $D \subset\subset \mathbb{R}^d$  and let  $u \in L^2(\Omega, L^2(D))$  be a zero-mean random field with a continuous covariance function, then*

$$u(\mathbf{x}, \omega) = \sum_{n \in \mathbb{N}} \sqrt{\lambda_n} z_n(\omega) \phi_n(\mathbf{x}), \quad (2.36)$$

where  $\{(\lambda_i, \phi_i)\}_{i \in \mathbb{N}}$  are the eigenvalue-eigenfunction pairs of the associated covariance operator  $\mathcal{C}$  and the series converges in  $L^2(\Omega, \mathbb{R})$  uniformly in  $\mathbf{x}$ . The  $z_n \in L^2(\Omega, \mathbb{R})$  are given by

$$\sqrt{\lambda_n} z_n = (u, \phi_i), \quad (2.37)$$

and are zero-mean random variables such that  $\mathbb{E}[z_i z_j] = \delta_{ij}$ .

*Remark 2.10.* In the case in which the field  $u$  is Gaussian, then the  $z_i$  are independent identically distributed (i.i.d.) standard Gaussian random variables.

The series is called the Karhunen-Loëve (KL) expansion of  $u$  and in practice it must be truncated after  $m_{\text{KL}}$  terms, where  $m_{\text{KL}}$  is the number of terms needed to achieve a given tolerance in the  $L^2(\Omega, L^2(D))$  norm. This depends on the rate of decay of the eigenvalues of  $\mathcal{C}$ , which in turn depends on the smoothness of  $C(\mathbf{x}, \mathbf{y})$  [151].

For practical computations the eigenfunctions of  $\mathcal{C}$  are often not known in advance and they must be estimated after discretisation by solving a (possibly generalised) dense eigenvalue problem (see e.g. [102] for an application with the FEM) which has a  $O(m_{\text{KL}}^2 m)$  computational cost and  $O(m_{\text{KL}} m)$  storage requirement, where  $m$  is the number of locations at which point values of  $u$  are needed and/or the dimension of the finite-dimensional approximation subspace used to approximate the  $\phi_i$ . If the eigenpairs are known *a priori* or if the  $\phi_i$  are approximated with compactly supported basis functions the cost per random field sample is  $O(m_{\text{KL}} m)$ .

The KL expansion method can be seen as the continuous equivalent of the factorisation methods just presented with the same upsides (flexible approach, works any point collection and almost any covariance function) and downsides (expensive if  $C(\mathbf{x}, \mathbf{y})$  is non-smooth). Additionally, the KL expansion provides a sampling method for non-Gaussian random fields as long as the  $z_i$  random variables can still be computed efficiently (in the general case, they are non-Gaussian and only uncorrelated, not necessarily independent).

### 2.4.3 Circulant embedding

The circulant embedding method<sup>2</sup> was first developed by Wood and Chan [176] and Dietrich and Newsam [49] and recently analysed by Graham et al. [73] and by

---

<sup>2</sup>The description that follows is adapted from [49] and [73].

Bachmayr et al. [13]. It computes samples of a stationary Gaussian field (cf. Definition 2.2)  $u$  evaluated at equispaced points on a structured grid and relies on the fast Fourier transform (FFT). The circulant embedding method works by the same principle exploited by equation (2.32). However, instead of factorising the covariance matrix  $C$  directly, it firsts embeds  $C$ , which on a structured uniform grid is block Toeplitz up to index reordering, into a bigger  $s$ -by- $s$  nested block circulant matrix  $Q$ . The eigen-decomposition of  $Q$ ,  $Q = G\Lambda G^T$  can be efficiently computed via the  $d$ -dimensional FFT in  $O(s \log s)$  complexity. Once the eigen-decomposition is available, equation (2.32) can be used with  $H = G\Lambda^{1/2}$ . The size of the embedding matrix  $s$  must be large enough to ensure positive definiteness as the block circulant extension is not necessarily positive definite [73].

We now give a brief description of the circulant embedding method for isotropic covariances (cf. Definition 2.2). For simplicity, we only consider the sampling over the unit interval in 1D.

The circulant embedding algorithm computes sample values of the Gaussian field at  $m_0 + 1$  equispaced points  $x_1, \dots, x_{m_0+1} \in [0, 1]$ , where  $x_i = i/(m_0 + 1)$ . The method requires the extension of the computational domain from  $[0, 1]$  to  $[0, l]$ , where  $l = m/m_0 > 1$ , with  $m > m_0$  integer. Let  $R$  be the covariance matrix over the extended domain at the points  $x_1, \dots, x_{m+1}$ , where  $x_i = i/(m + 1)$  and the first  $m_0 + 1$  points coincide with the previous points. Then  $R$  with entries  $R_{ij} = C(|x_i - x_j|)$  is a positive semi-definite Toeplitz matrix and as such is uniquely determined by its first row  $\mathbf{r}$ , where  $r_i = C(|x_1 - x_i|)$ . This matrix can be embedded within a  $s \times s$  Toeplitz matrix  $Q$ , with  $s = 2(m + 1)$  and whose first row  $\mathbf{q}$  is given entrywise by

$$q_k = r_k, \quad 1 \leq k \leq m + 1, \tag{2.38}$$

$$q_{2(m+1)-k} = r_k, \quad 0 \leq k \leq m. \tag{2.39}$$

$Q$  is then symmetric and circulant, and admits the eigen-decomposition  $Q = G\Lambda G^T$ , where  $G = \text{Re}(F) + \text{Im}(F)$  and  $F$  is the FFT matrix of size  $s$  with entries  $F_{jk} = s^{-1/2} \exp(2\pi i j k / s)$  (here  $i = \sqrt{-1}$ ) (lemma 4 in [76]) and  $\Lambda$  is the diagonal matrix of the eigenvalues of  $Q$ , which are given by  $\boldsymbol{\lambda} = F\mathbf{q}$  [49, 73].

If the extension parameter  $l$  is large enough, then all entries in  $\Lambda$  are positive and a sample from the Gaussian field  $u$  can be computed as follows (taken from [73]):

1. Sample an  $s$ -dimensional standard Gaussian vector  $\mathbf{z} \sim \mathcal{N}(0, I)$ .
2. Compute  $\boldsymbol{\lambda} = F\mathbf{q}$  by applying the FFT to  $\mathbf{q}$ . This step can be done offline.
3. Compute  $\mathbf{w} = F(\sqrt{\boldsymbol{\lambda}} \odot \mathbf{z})$ , where  $\odot$  indicates the Hadamard entrywise product and the action of  $F$  can be computed efficiently via the FFT.
4. Set  $\mathbf{u}^{\text{ext}} = \text{Re}(\mathbf{w}) + \text{Im}(\mathbf{w})$ . The vector  $\mathbf{u}^{\text{ext}}$  contains the  $m_0 + 1$  point values of  $u$  needed.

The overall cost complexity of the circulant embedding algorithm is  $O(s \log s)$  with an  $O(s)$  memory requirement, where  $s$  depends on the extension factor  $l$  which must

be large enough so that all the eigenvalues of  $Q$  are positive. Graham et al. in [73] derive conditions on  $C(\mathbf{x}, \mathbf{y})$  and on  $l$  such that this happens in general dimensions. In the Matérn case, these conditions were recently improved by Bachmayr et al. in [13] and read:

**Theorem 2.5** (Theorem 10 in [13]). *Consider the Matérn covariance family (2.27) with smoothness parameter  $0 < \nu < \infty$  and correlation length  $\lambda > 0$ . Then there exist positive constants  $C_1$  and  $C_2$  dependent on  $d$  only such that  $Q$  is positive definite if*

$$l/\lambda \geq C_1 + C_2 \max(\nu^{1/2}(1 + |\log \nu|), \nu^{-1/2}). \quad (2.40)$$

Overall, the circulant embedding is an extremely fast algorithm and works for quite a general class of covariance matrices provided that a finite extension factor  $l$  can be found [73]. The downside of this method is that its efficiency depends on the use of simple geometries. Once the field is sampled, it must be transferred to the target mesh on which the field is required. If the sampling and target meshes are non-nested, then interpolation or projection is needed. In [74], the authors prove that if the uniform grid size is asymptotically equivalent to the maximum mesh size of the target mesh and the sampled field is  $C^1(D)$  or less, then multilinear interpolation can be used without harming the overall FEM convergence order. However, for smoother fields, we suspect that a more careful treatment would be required to prevent the low smoothness ( $W^{1,\infty}(D)$ ) of the interpolated field to harm the convergence of the quadrature rule (e.g. a supermesh projection could be used, cf. Chapter 4).

#### 2.4.4 The SPDE approach to Matérn field sampling

We conclude this section by describing the finite element basis approach, called the *SPDE approach to Matérn field sampling*. This is the only method we use in this thesis.

Whittle showed in [138] that a Matérn field with covariance given by (2.27) is the statistically stationary solution that satisfies the linear elliptic PDE,

$$(\mathcal{I} - \kappa^{-2} \Delta)^k u(\mathbf{x}, \omega) = \eta \dot{\mathbf{W}}, \quad \mathbf{x} \in \mathbb{R}^d, \quad \omega \in \Omega, \quad \nu = 2k - d/2 > 0, \quad (2.41)$$

where  $\dot{\mathbf{W}}$  is spatial Gaussian white noise in  $\mathbb{R}^d$  (cf. Definition 2.6),  $k > d/4$ . Here  $d \leq 3$  and the equality has to hold almost surely and be interpreted in the sense of distributions. Boundary conditions are not needed as the stationarity requirement is enough for well-posedness [116]. The constant  $\eta$  is a scaling factor that depends on  $\sigma$ ,  $\lambda$  and  $\nu$ ,

$$\eta = \frac{\sigma}{\hat{\sigma}}, \quad \text{where} \quad \hat{\sigma}^2 = \frac{\Gamma(\nu)}{\Gamma(\nu + d/2)} \left( \frac{2}{\pi} \right)^{d/2} \lambda^{-d}, \quad (2.42)$$

where  $\Gamma(x)$  is the Euler gamma function [116]. Note that if  $d = 2$  then  $\hat{\sigma}^2 = (2/\pi)\lambda^{-2}$ , and for  $\nu \rightarrow \infty$ ,  $\hat{\sigma}^2 = (2/\pi)^{d/2}\lambda^{-d}$ . In this thesis we assume the value of  $k$  in equation (2.41) to be a positive integer, although it is possible to work with non-integer values as well [29, 42].

Equation (2.41) has to be solved on the whole  $\mathbb{R}^d$ . However this is generally not feasible and  $\mathbb{R}^d$  is in practice truncated to a bounded domain  $D \subset\subset \mathbb{R}^d$ . In this case, artificial boundary conditions must be prescribed on  $\partial D$ . Homogeneous Dirichlet or Neumann boundary conditions are often chosen in the literature [27, 116], although it usually does not matter for practical purposes as the error in the covariance of the field decays rapidly away from the boundary (see [141] for a practical study and [103] for the theoretical analysis). In a recent study, Khristenko et al. [103] show that homogeneous Robin boundary conditions work better in practice and are exact for the 1D case. In what follows we assume that the Matérn field sample is needed on a domain  $G \subset\subset D$  and we prescribe homogeneous Dirichlet boundary conditions on  $\partial D$ . If  $D$  is sufficiently large in the sense that the distance between  $\partial D$  and  $\partial G$  is large enough, then the error introduced by truncating  $\mathbb{R}^d$  to  $D$  is negligible (cf. Remark 2.12 and [51, 103, 141]).

Once we truncate  $\mathbb{R}^d$  to  $D$  and apply Dirichlet-type boundary conditions, (2.41) can be rewritten in the following iterative form,

$$\begin{cases} u_1 - \kappa^{-2} \Delta u_1 = \eta \dot{W} & \text{in } D \\ u_{j+1} - \kappa^{-2} \Delta u_{j+1} = u_j & \text{in } D, \quad j = 1, \dots, k-1, \\ u_{j+1} = 0 & \text{on } \partial D, \quad j = 0, \dots, k-1, \end{cases} \quad (2.43)$$

where  $u \equiv u_k$ . This is the approach suggested by Lindgren et al. in [116] and requires solving a sequence of second order elliptic problems to produce a sample of  $u$ . In the  $k=1$  case (2.43) reduces to

$$\begin{aligned} u - \kappa^{-2} \Delta u &= \eta \dot{W} && \text{in } D, \\ u &= 0 && \text{on } \partial D. \end{aligned} \quad (2.44)$$

Existence and uniqueness of an a.s.  $C^0(\bar{D})$  solution to (2.44) was proven in [20] and in [32], from which existence and uniqueness of the iterative problem (2.43) follows from standard theory.

After the meshing of  $D$ , (2.41) and (2.43) can be solved in linear time with the finite element method (FEM) and an optimal solver (e.g. full multigrid). This sampling approach thus scales well in terms of problem size and parallel computation provided  $k$  is not excessively large<sup>3</sup> [51]. Moreover, the approach is especially convenient if  $u$  appears as a coefficient in a PDE which is solved using the FEM as it might be possible to reuse finite element bases and computations for both equations.

Let  $s$  be the number of degrees of freedom of the finite element subspace used to solve (2.41). The memory and set-up cost complexity are in both cases  $O(s)$ , while the cost of computing one sample  $u$  with the SPDE approach is  $O(ks)$  provided that we can pick  $k$  in (2.41) to be an integer and that we are using an optimal solver such as a full multigrid method. The cost is  $O(ks + s \log s)$  otherwise as a fractional order PDE must be solved for each sample [27]. All these considerations assume that white noise can also be sampled in  $O(s)$  complexity. Showing how this can be done

---

<sup>3</sup>If  $k$  is large, then  $\nu$  is large, hence the field is smooth and the other sampling methods we presented are more efficient.

in practice in a non-nested MLMC and MLQMC framework is one of the main topics of this thesis and we refer to chapters 4 and 5 for the details.

*Remark 2.11.* In principle, if the values of  $u$  are only needed on a structured grid, one could use FFT techniques to solve (2.41). However, in a general setting it might not be possible to use the FFT as some applications might require either unstructured grids or a generalised version of (2.41) with variable coefficients [116].

*Remark 2.12.* The error introduced by solving (2.41) on a truncated domain decreases exponentially as the distance  $\delta$  from the boundary increases with a decay proportional to  $C(r)|_{r=2\kappa\delta} = C(2\kappa\delta)$ , where  $C$  is the Matérn covariance function of the sampled field and  $\kappa$  depends on the Matérn parameters (cf. equation (2.27)) [103]. For this reason, we assume that the domain  $D$  is always large enough so that this error is negligible in the inner domain  $G$  where the Matérn field sample is needed. In this thesis we always choose  $D$  to be a cube although other choices are possible.

# Chapter 3

## Finite element convergence analysis for spatial white noise driven elliptic PDEs

In this chapter we address the proof of convergence for the FEM solution of problem (2.43) and derive *a priori* FEM strong error estimates in the  $h$ - and  $p$ -refinement cases, i.e. when either the polynomial degree of the FEM basis is fixed and the mesh is refined or the mesh is fixed and the polynomial degree is increased. In the first and second section we address the proof of the case  $k = 1$  and in the third section we deal with the  $k > 1$  case.

*Remark 3.1* (On the proof novelty and related work). The proof presented in this chapter was derived by the author in the period between January and April of 2017. At that time, the main result of this proof, namely that there is no need to approximate white noise in practice to ensure convergence of the FEM, was novel (the convergence rates for  $h$ -refinement were not [36, 52, 157]). Although the structure of the proof is still new nowadays, a more general convergence proof yielding a more powerful result was concurrently developed by Bolin et al. in [27], originally uploaded to the arXiv in May 2017. Among other things, Bolin et al. consider the more general fractional case and obtain the optimal convergence order of  $O(h^{\min(\nu,p+1)})$ , where  $\nu$  is the smoothness parameter of the Matérn field and  $p$  is the polynomial degree of the FEM basis. Our proof, unlike that of Bolin et al., also considers a  $p$ -refinement approach in which convergence is achieved by increasing the polynomial degree of the FEM basis on a fixed mesh. However, as we will see in this chapter, we were only able to obtain a rate of  $O(p^{-(\nu-\epsilon)}h^{\min(\nu-\epsilon,p+1)})$  for any  $\epsilon > 0$  arbitrarily small. It is also worth mentioning some other more recent related work on the subject: in [28] Bolin et al. derive *a priori* weak error estimates for the solution of (2.41) on a bounded domain and in [42] Cox and Kirchner derive *a priori* strong error estimates and regularity results in Sobolev and Hölder spaces for the solution of the same equation and its covariance function.

The main challenge in establishing a convergence result for equation (2.43) is the lack of smoothness of white noise. Realisations of white noise are a.s. in  $H^{-d/2-\epsilon}(D)$

for all  $\epsilon > 0$  (cf. Remark 2.8) and therefore cannot be paired with e.g. Lagrange basis functions which are only  $W^{1,\infty}(D)$  (at least for  $d = 3$ ). For this reason, existing convergence proofs approximate white noise with a sequence of smooth functions that converges to white noise in the limit. Common choices in the literature are either a piecewise constant approximation [36, 52, 157] or a truncated spectral series expansion [182]. Another option, not investigated in the references just mentioned, is to avoid sampling white noise realisations directly and to sample its action onto a test function instead. This idea is at the heart of the convergence proof presented in this chapter. In fact, white noise is also in  $\mathcal{L}(L^2(D), L^2(\Omega, \mathbb{R}))$ , i.e. it is a linear functional from  $L^2(D)$  to the space of random variables with finite variance. Therefore its action onto a test function  $v \in L^2(D)$ , namely  $\langle \dot{W}, v \rangle(\omega)$ , is a Gaussian random variable which is a.s. finite and can easily be sampled. Lagrange basis functions are all in  $L^2(D)$  and can all therefore be paired with white noise, provided that we sample  $\langle \dot{W}, v \rangle(\omega)$  rather than  $\dot{W}(\cdot, \omega)$  (recall Remark 2.7).

The advantage of doing this is that the white noise action can be computed exactly, without the need for approximations. As we will show in this chapter, it is possible to work with exact white noise and still achieve optimal FEM convergence. Furthermore, sampling the action fits well into the MLMC and MLQMC framework we will be presenting in chapters 4 and 5 as it simplifies the coupling of white noise on each level.

### 3.1 Finite element convergence for problem (2.43): the case $k = 1$

We want to investigate the FEM solution of problem (2.43). For simplicity, we set  $\eta = 1$  as this does not impact the generality of the result. We start by considering the case of the second order linear elliptic SPDE corresponding with the case  $k = 1$ ,

$$u - \kappa^{-2} \Delta u = \dot{W}, \quad \text{in } D, \quad u = 0 \quad \text{in } \partial D. \quad (3.1)$$

The solution to this problem exists and is unique in  $L^q(\Omega, L^2(D))$  for all  $q \in [1, \infty)$  as proven in [32]. It also has the following regularity properties,

**Theorem 3.1** ([32, 42, 182]). *Assuming  $D$  is a bounded Lipschitz domain and  $d = 1, 2, 3$ , the (very) weak solution to equation (3.1) exists and is unique in  $L^q(\Omega, L^2(D))$  for all  $q$  and it satisfies*

$$u \in H^{2-d/2-\epsilon}(D) \cap C_0^{\lfloor 2-d/2-\epsilon \rfloor}(\bar{D}), \quad \text{a.s.}, \quad (3.2)$$

*As  $u$  is a.s. continuous on a compact set, it is also in  $L^q(D)$  for  $q \in [1, \infty]$  a.s.*

The main focus of the following sections is to prove convergence for the finite element approximation of the solution of equation (3.1) with standard Lagrange elements. We simplify the exposition by introducing the following assumptions.

### 3.1.1 Assumptions on the domain $D$

As previously stated in observation 2.12, we will consider the numerical solution of (2.43) on a bounded domain  $D$  which is large enough to make the error introduced by restricting  $\mathbb{R}^d$  to  $D$  negligible in the inner region  $G$ .

We are totally free to choose the shape of the domain  $D$ . For theoretical purposes, choosing  $D$  to be a smooth domain (i.e. a sphere) makes the convergence analysis easier as this choice yields solutions of (2.43) which are as smooth as they can be. For practical purposes, however, curved domain boundaries are difficult to work with and it is much more convenient to choose  $D$  to be a cuboidal domain. This choice, however, has the disadvantage of affecting the smoothness of the solution for large  $k$  [78], which would then affect the FEM convergence.

In practice, the fact that we are only interested in the FEM error over  $G$ , which is in the *interior* of the domain, allows us to get the best of both worlds. Assume that there exists a subdomain<sup>1</sup>  $D_0$  of  $D$  such that  $\bar{G} \subset D_0 \subset\subset D$  (see Figure 3.1), where  $\bar{G}$  is the closure of  $G$ . Standard elliptic regularity theory [56] tells us that the lack of smoothness of  $\partial D$  only affects the solution regularity on  $\bar{D}$  and not on a compact subset such as  $\bar{G}$ , i.e. the smoothness of the solution in the interior is not capped by the smoothness of  $\partial D$ .

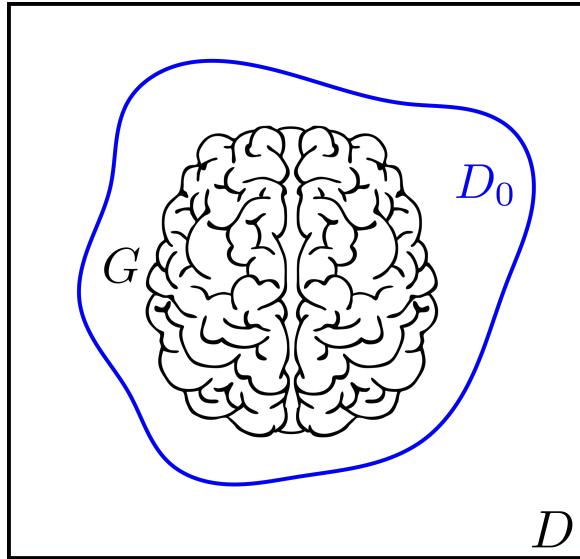


Figure 3.1: An example of a domain  $D_0$ , chosen arbitrarily such that  $\bar{G} \subset D_0 \subset\subset D$  (we just need such a  $D_0$  to exist), in the case in which  $D$  is a box domain and  $G$  is a brain domain. Although we sample the Matérn field on a large domain  $D$ , we are only interested in its values in the interior domain  $G$ .

Nevertheless, typical FEM interior error estimates (see [129]) tell us that the accuracy of the FEM in  $G$  still decreases as a result of the pollution effect arising from the non-smoothness of  $\partial D$ . It turns out that in the very specific case of  $h$ - and  $hp$ -refinement and  $D$  being a cuboidal domain, this pollution error is negligible and

---

<sup>1</sup>This assumption is technical, see for example the work by Nitsche and Schatz [129].

the same convergence estimates as in the  $C^\infty$  boundary case can be obtained when considering the FEM error over  $G$  (see [129]) at the cost of a slightly more complicated convergence analysis. In this thesis we use a box domain in practice, but, for the sake of simplicity, we assume that  $D$  has a  $C^\infty$  boundary in the FEM convergence analysis and we will consider the FEM approximation error over  $D$  instead of over  $G$ .

*Remark 3.2.* In the  $p$ -refinement case, the situation is different and the pollution error arising from a non-smooth boundary is not negligible, even in the case of a box domain [12]. In practice, Babuška and Suri observe in [12] that the pollution error arises if there are not enough mesh cells separating the inner region  $G$  from the points in which  $\partial D$  is non-smooth. This effect becomes more prominent and always appears as the polynomial degree gets larger. In this chapter, we will make the additional simplifying assumption that there are always enough mesh cells separating  $G$  from  $D$  so that the boundary pollution error is negligible for the range of polynomial degrees considered. Even in the  $p$ -refinement case, we will therefore simplify the analysis by considering  $D$  to have a  $C^\infty$  boundary. In this thesis we mainly use  $h$ -refinement and  $p$ -refinement is only used in Chapter 4.

Before proving convergence for the finite element approximation of the solution of equation (3.1), we give a brief review of the approaches used in the literature to prove similar results. We will be using these as a starting point in our proof.

### 3.1.2 Known techniques: introducing an auxiliary problem

The weak form of equation (3.1) reads: find  $u \in U$  such that

$$(u, v) + \kappa^{-2}(\nabla u, \nabla v) = \langle \dot{W}, v \rangle, \quad \text{a.s. } \forall v \in V, \quad (3.3)$$

where  $U$  and  $V$  are suitable Banach spaces. In this chapter we prove that taking  $U \equiv V \equiv H_0^1(D)$  is sufficient for convergence. Consider a quasi-uniform triangulation of  $D$  of maximum cell size  $h$  and let  $U_h \subseteq U$  and  $V_h \subseteq V$  be finite element approximation subspaces. Let  $u_h \in U_h$  be the FEM approximation of  $u$ . The discrete weak form is: find  $u_h \in U_h$  such that

$$(u_h, v_h) + \kappa^{-2}(\nabla u_h, \nabla v_h) = \langle \dot{W}, v_h \rangle, \quad \text{a.s. } \forall v_h \in V_h. \quad (3.4)$$

The poor regularity of white noise makes a standard finite element approach difficult for  $d \in \{2, 3\}$ . Since  $\dot{W} \in H^{-d/2-\epsilon}(D)$  a.s., we would need  $V \subseteq H_0^{d/2+\epsilon}(D)$  for the weak form to make sense. However, for Lagrange finite elements, we might have  $V_h \not\subseteq H_0^{d/2+\epsilon}(D)$  and smoother elements would be needed (such as Argyris, see [31]).

This means that we cannot immediately apply Lagrange elements to solve the weak form of equation (3.1) directly. However, inspired from previous work on the subject [36, 52, 182], we define the auxiliary problem

$$u^m - \kappa^{-2}\Delta u^m = \dot{W}^m, \quad \text{in } D, \quad u^m = 0 \quad \text{in } \partial D, \quad (3.5)$$

where now  $\dot{W}^m$  is a smooth approximation of the white noise  $\dot{W}$  (in the literature  $\dot{W}^m \in L^2(D)$  a.s. for finite  $m$ ) such that  $\dot{W}^m \rightarrow \dot{W}$  as  $m \rightarrow \infty$  in some sense, where

$m$  is the accuracy of the approximation. We fix a relation between  $m$ , the cell size  $h$  and the polynomial degree  $p$  of the Lagrange elements, so that  $m = m(h, p)$  and  $m(h, p) \rightarrow \infty$  as  $h/p \rightarrow 0$ .

Let  $u_h^m$  be the finite element approximation to the solution of (3.5). As  $\dot{W}^m$  is smooth, the weak form of (3.5) reads: find  $u^m \in H_0^1(D)$  such that

$$(u^m, v) + \kappa^{-2}(\nabla u^m, \nabla v) = (\dot{W}^m, v), \quad \text{a.s. } \forall v \in H_0^1(D), \quad (3.6)$$

where we do not need the angle bracket notation since  $\dot{W}^m$  is in  $L^2(D)$  a.s. (see Lemma 3.5 in the next section for a proof). Let  $V_h$  be the FEM approximation subspace spanned by standard Lagrange elements and let  $u_h^m \in V_h$  be the FEM approximation to  $u^m$ . Then  $u_h^m$  satisfies,

$$(u_h^m, v_h) + \kappa^{-2}(\nabla u_h^m, \nabla v_h) = (\dot{W}_h^m, v_h), \quad \text{a.s. } \forall v \in V_h. \quad (3.7)$$

The approach used in [36, 52, 182] is to show that  $u_h^m \rightarrow u$  as  $h \rightarrow 0$ . The idea is to bound the error between  $u$  (exact solution of (3.1)) and  $u_h^m$  by

$$\mathbb{E}[||u - u_h^m||_{L^2(D)}^2] = \mathbb{E}[||u - u^m + u^m - u_h^m||_{L^2(D)}^2] \quad (3.8)$$

$$\leq 2\mathbb{E}[||u - u^m||_{L^2(D)}^2] + 2\mathbb{E}[||u^m - u_h^m||_{L^2(D)}^2], \quad (3.9)$$

where we used the linearity of the expected value and the triangle inequality.

If the approximation  $\dot{W}^m$  is chosen so that  $\dot{W}^m$  is smooth enough (i.e.  $\dot{W}^m \in L^2(D)$  a.s.), then equation (3.5) can be solved with Lagrange elements and convergence to zero for the second term in (3.9) can be shown by using standard techniques (more or less, see later in the section). If  $\dot{W}^m$  can also be chosen so that  $u^m$  converges to  $u$  fast enough as  $m \rightarrow \infty$ , then we get convergence for the total error  $\mathbb{E}[||u - u_h^m||_{L^2(D)}^2]$ .

### 3.1.3 The main result

The main focus of the next sections is the proof of the following theorem.

**Theorem 3.2.** *Consider a quasi-uniform triangulation of  $D \subset \subset \mathbb{R}^d$  of maximum cell size  $h$ . Let  $u \in H^s$ , with  $s = 2 - d/2 - \epsilon$  be the solution to (3.1) and let  $u_h \in V_h$  be the finite element approximation of  $u$  obtained using Lagrange elements of degree  $p \geq 1$  on each cell, then, for all sufficiently small  $\epsilon > 0$  and for either fixed  $p$  or fixed  $h$ , there exists a  $\hat{C}(\epsilon, s, d, D) \in \mathbb{R}$  blowing up as  $\epsilon \rightarrow 0$  such that*

$$\mathbb{E}[||u - u_h||_{L^2(D)}^2]^{1/2} \leq \hat{C}(\epsilon, s, d, D)h^{2-d/2-\epsilon}p^{-(2-d/2-\epsilon)}. \quad (3.10)$$

This is the first time a  $p$ -refinement estimate has been provided for such a problem. However, the main novelty of Theorem 3.2 is not in the order of convergence of the FEM approximation: an  $h$ -refinement estimate of the same convergence order as (3.10) has been proven in many papers in the literature (see for Example [36, 52, 182] and many more). In all these papers, however, the authors do not compute the white noise exactly, but they work with an approximated white noise. In this thesis,

we show that in practice approximating the white noise is not needed and that it is possible to show that the same order convergence is achievable by using Lagrange elements as they are. A similar result with an improved convergence order without the  $\epsilon$  term was also derived by Bolin et al. [27] (cf. Remark 3.1).

Typical approximations used in previous work were either a truncated spectral expansion [182] or a piecewise constant approximation of white noise [36, 52]. We use a similar approach as in [36, 52, 182] to prove convergence, but we use an “approximation” of white noise which is exact if applied to any function in the finite element approximation subspace  $V_h$ , i.e.  $(\dot{W}^m, v_h) = \langle \dot{W}, v_h \rangle$ , for all  $v_h \in V_h$ . As in the FEM we only use test functions that belong to  $V_h$ , no approximation is actually needed in practice. This also means that  $U_h \equiv V_h$  and  $u_h^m \equiv u_h$  in our case and this is why  $u_h^m$  does not appear in Theorem 3.2.

## 3.2 Proof of Theorem 3.2

### 3.2.1 Our “approximation” of white noise

In this section, we show that we can avoid approximating the white noise  $\dot{W}$  in practice by using the approximate  $\dot{W}^m$  exclusively as a theoretical tool. To do this, we first need a tool commonly used in finite element practice: the orthogonal  $L^2$ -projection onto  $V_h$ .

**Definition 3.1** (orthogonal projection onto  $V_h$ ). Let  $v \in L^2(D)$  and let  $V_h$  be a finite element approximation space. Then  $P_h : L^2(D) \rightarrow V_h$  such that

$$P_h v = \arg \min_{w_h \in V_h} \|v - w_h\|_{L^2(D)} \quad (3.11)$$

is the orthogonal  $L^2$ -projection onto the finite element space  $V_h$ .

**Theorem 3.3** (approximation properties of  $V_h$  [12, 21, 22, 152]). *Let  $V_h$  be a FEM subspace spanned by Lagrange elements of degree  $p \geq 1$  on each cell of a quasi-uniform triangulation of  $D$  of maximum mesh size  $h$ . Let  $v \in H^s(D)$ ,  $s > 0$ , then for either fixed  $p$  or fixed  $h$  the best approximation of  $v$  in  $V_h$  satisfies,*

$$\min_{w_h \in V_h} \|v - w_h\|_{L^2(D)} \leq c(s, d, D) h^\mu p^{-s} \|v\|_{H^s(D)}, \quad \mu = \min(s, p + 1). \quad (3.12)$$

Note that  $c$  is allowed to depend on  $s$ ,  $D$ ,  $d$ , but is independent from  $v$ ,  $p$  or  $h$ .

*Proof.* For fixed  $h$  see the work by Babuška and Suri [12] and by Schwab [152]. For fixed  $p$ , note that the best approximation error is lesser or equal than the global interpolation error of  $v$  in  $V_h$ . Choosing the Clemént interpolant (see [21, 22]) yields the bound.  $\square$

*Remark 3.3.* Note that, by definition, we have  $P_h v_h = v_h$ ,  $\forall v_h \in V_h$ .

We can now define our approximation to white noise:

**Definition 3.2** (projected white noise). Let  $V_h$  be the Lagrange finite element approximation subspace used to discretise (3.3) and let  $P_h : L^2(D) \rightarrow V_h$  be the orthogonal projection onto  $V_h$ . Let  $m$  be the number of basis functions that span  $V_h$ , then we define the projected white noise as the generalised stochastic field given by,

$$(\dot{W}^m, v) := \dot{W}(P_h v) \equiv \langle \dot{W}, P_h v \rangle, \quad \text{i.e.} \quad \dot{W}^m := \dot{W} \circ P_h, \quad \forall v \in L^2(D). \quad (3.13)$$

*Remark 3.4.* The reason why we claim that  $\dot{W}^m$  is just a theoretical tool and does not change anything from a practical aspect is that, with this choice of  $\dot{W}^m$ , we have, for all  $v_h \in V_h$ ,

$$(\dot{W}^m, v_h) = \langle \dot{W}, P_h v_h \rangle = \langle \dot{W}, v_h \rangle, \quad \forall v_h \in V_h. \quad (3.14)$$

This means that  $\dot{W}^m$  coincides with  $\dot{W}$  in  $V_h$ . Hence, the finite element discretisations of (3.1) and of (3.5) are exactly the same and  $u_h \equiv u_h^m$ , where  $u_h$  is the finite element approximation of the solution of (3.1).

As we said before, we require two properties from  $\dot{W}^m$ :

1.  $\dot{W}^m$  must converge to  $\dot{W}$  as  $m \rightarrow \infty$  in some sense.
2.  $\dot{W}^m$  must be in  $L^2(D)$  a.s. for finite  $m$ .

These two properties are proved in the following two subsections.

### $\dot{W}^m$ converges to $\dot{W}$ as $m \rightarrow \infty$ in $\mathcal{L}(\mathbf{H}^s(\mathbf{D}), \mathbf{L}^2(\Omega, \mathbb{R}))$ for $s > 0$

The first property is given by the following lemma,

**Lemma 3.4.** *Let  $\dot{W}^m$  be the projected white noise as defined in Definition 3.2. Let  $V_h \subseteq L^2(D)$  be the Lagrange finite element approximation subspace of degree  $p$  used to discretise (3.1). Let  $s > 0$ , let  $v \in H^s(D)$  and  $\mu = \min(s, p + 1)$ . Then for one between  $p$  or  $h$  fixed the following bounds hold:*

$$\mathbb{E}[|\langle \dot{W} - \dot{W}^m, v \rangle|^2]^{1/2} \leq c(s, d, D)p^{-s}h^\mu \|v\|_{H^s(D)}, \quad (3.15)$$

$$\|\dot{W} - \dot{W}^m\|_{\mathcal{L}(H^s(D), L^2(\Omega, \mathbb{R}))} \leq c(s, d, D)p^{-s}h^\mu. \quad (3.16)$$

*Proof.* We have that, for all  $v \in L^2(D)$ ,

$$\mathbb{E}[|\langle \dot{W} - \dot{W}^m, v \rangle|^2] = \mathbb{E}[|\langle \dot{W}, v \rangle - \langle \dot{W}, P_h v \rangle|^2] \quad (3.17)$$

$$= \mathbb{E}[|\langle \dot{W}, v - P_h v \rangle|^2] = \|v - P_h v\|_{L^2(D)}^2, \quad (3.18)$$

by the definition of  $\dot{W}$  and the definition of the projected white noise  $\dot{W}^m$ . As  $P_h v$  is the best approximation to  $v$  in the  $L^2(D)$  norm in the finite element approximation space  $V_h$  the first bound follows from Theorem 3.3.

For the second bound, Theorem 3.3 gives,

$$\|\dot{W} - \dot{W}^m\|_{\mathcal{L}(H^s(D), L^2(\Omega, \mathbb{R}))} = \sup_{\|v\|_{H^s(D)} \leq 1} \mathbb{E}[|\langle \dot{W} - \dot{W}^m, v \rangle|^2]^{1/2} \quad (3.19)$$

$$= \sup_{\|v\|_{H^s(D)} \leq 1} \|v - P_h v\|_{L^2(D)} \leq c(s, d, D) p^{-s} h^\mu \sup_{\|v\|_{H^s(D)} \leq 1} \|v\|_{H^s(D)} \leq c(s, d, D) p^{-s} h^\mu, \quad (3.20)$$

where the suprema are taken over all  $v \in H^s(D)$ .  $\square$

This means that the action of  $\dot{W}^m$  onto smooth enough test functions (at least in  $H^s(D)$  for  $s > 0$ ) converges to the action of  $\dot{W}$  as we refine the mesh or as we increase the polynomial degree.

### $\dot{W}^m$ is in $L^2(\mathbf{D})$ a.s. for finite $m$

The second property is given by the following lemma,

**Lemma 3.5.** *Let  $\dot{W}^m$  be the projected white noise as defined in Definition 3.2. Let  $V_h \subseteq L^2(D)$  be the finite element approximation space used to discretise (3.1). Let  $m$  be the number of basis functions that span  $V_h$ . Then,  $\dot{W}^m \in L^2(\Omega, L^2(D))$  for finite  $m$  and for all  $q \in (0, \infty)$ ,*

$$\mathbb{E}[\|\dot{W}^m\|_{L^2(D)}^q]^{1/q} = \varphi(m, q)\sqrt{m}, \quad \text{where } \varphi(m, q) = \left( \frac{\Gamma(m/2 + q/2)}{\Gamma(m/2)(m/2)^{q/2}} \right)^{1/q}. \quad (3.21)$$

The function  $\varphi(m, q)$  satisfies the following properties: for all fixed  $q \in (0, \infty)$ ,  $\varphi(m, q) \sim 1$  as  $m \rightarrow \infty$ ; for  $q \leq 2$ ,  $\varphi(m, q) \leq 1$  for all  $m$  and  $\varphi(m, 2) \equiv 1$ . We therefore have

$$\mathbb{E}[\|\dot{W}^m\|_{L^2(D)}^2]^{1/2} = \sqrt{m}. \quad (3.22)$$

*Proof.* We know that

$$\|\dot{W}^m\|_{L^2(D)} = \sup_{\|v\|_{L^2(D)} \leq 1} |\langle \dot{W}^m, v \rangle| = \sup_{\|v\|_{L^2(D)} \leq 1} |\langle \dot{W}, P_h v \rangle|, \quad (3.23)$$

by the properties of projected white noise (see Definition 3.2).

We know that the projection operator  $P_h$  is stable in  $L^2(D)$ , i.e.,

$$\|P_h v\|_{L^2(D)} \leq \|v\|_{L^2(D)}, \quad \forall v \in L^2(D), \quad (3.24)$$

see for Example [55]. By using stability, we have the following equivalence,

$$\{v_h \in V_h : \exists v \in L^2(D), \|v\|_{L^2(D)} \leq 1, v_h = P_h v\} \equiv \{v_h \in V_h : \|v_h\|_{L^2(D)} \leq 1\}. \quad (3.25)$$

Hence, equation (3.23) becomes

$$\|\dot{W}^m\|_{L^2(D)} = \sup_{\|v\|_{L^2(D)} \leq 1} |\langle \dot{W}, P_h v \rangle| = \sup_{\substack{v_h \in V_h \\ \|v_h\|_{L^2(D)} \leq 1}} |\langle \dot{W}, v_h \rangle|. \quad (3.26)$$

Now, let  $\{\phi_i\}$ ,  $i = 1, \dots, m$ , be the basis functions that span  $V_h$ . Without loss of generality, we can assume that they are orthonormal in  $L^2(D)$ . Therefore, for every  $v_h \in V_h$ , there exists a vector of coefficients  $\mathbf{a} \in \mathbb{R}^m$  such that  $v_h = \sum_i a_i \phi_i$ , with  $\|v_h\|_{L^2(D)} = |\mathbf{a}|_2$ , by Parseval's equality. Thanks to the definition of white noise, we obtain

$$\langle \dot{W}, v_h \rangle = \sum_i a_i \langle \dot{W}, \phi_i \rangle = \mathbf{a}^T \mathbf{z}, \quad \text{where } \mathbf{z} \sim \mathcal{N}(0, I), \quad \forall v_h \in V_h. \quad (3.27)$$

We then have

$$\|\dot{W}^m\|_{L^2(D)} = \sup_{\substack{v_h \in V_h \\ \|v_h\|_{L^2(D)} \leq 1}} |\langle \dot{W}, v_h \rangle| = \sup_{\substack{\mathbf{a} \in \mathbb{R}^m \\ |\mathbf{a}|_2 \leq 1}} |\mathbf{a}^T \mathbf{z}|, \quad \text{a.s.} \quad (3.28)$$

The inner product  $|\mathbf{a}^T \mathbf{z}|$  is maximised for fixed  $\mathbf{z}$  when  $\mathbf{a} = \mathbf{z}/|\mathbf{z}|_2$ , hence

$$\|\dot{W}^m\|_{L^2(D)} = \sup_{\substack{\mathbf{a} \in \mathbb{R}^m \\ |\mathbf{a}|_2 \leq 1}} |\mathbf{a}^T \mathbf{z}| = |\mathbf{z}|_2, \quad \text{a.s.} \quad (3.29)$$

The random variable  $|\mathbf{z}|_2$  is a chi random variable with  $m$  degrees of freedom and density  $2^{1-m/2} x^{m-1} \exp(-x^2/2)/\Gamma(m/2) \mathbb{1}_{\{x>0\}}(x)$ , where  $\Gamma(x)$  is the Euler Gamma function and  $\mathbb{1}_{\{x>0\}}(x)$  is the indicator function of the set  $\{x > 0\}$  [63]. We then conclude

$$\mathbb{E}[\|\dot{W}^m\|_{L^2(D)}^q]^{1/q} = \mathbb{E}[|\mathbf{z}|_2^q]^{1/q} = \left( \int_0^\infty \frac{x^{m-1+q} e^{-x^2/2}}{2^{m/2-1} \Gamma(m/2)} dx \right)^{1/q} = \sqrt{m} \varphi(m, q), \quad (3.30)$$

$$\text{where } \varphi(m, q) = \left( \frac{\Gamma(m/2 + q/2)}{\Gamma(m/2)(m/2)^{q/2}} \right)^{1/q}. \quad (3.31)$$

For fixed  $q$ ,  $\varphi(m, q) \sim 1$  as  $m \rightarrow \infty$ ; for  $q \in (0, 2]$ ,  $\varphi(m, q) \leq 1$  and  $\varphi(m, 2) \equiv 1$  [175]. Consequently  $\dot{W}^m \in L^2(\Omega, L^2(D))$  for finite  $m$  and therefore  $\dot{W}^m \in L^2(D)$  a.s. for finite  $m$ , and this concludes the proof.  $\square$

*Remark 3.5* (Why is this result important?). The fact that  $\dot{W}^m \in L^2(D)$  a.s. for finite  $m$  is extremely important as this implies, under suitable conditions, that  $u^m \in H^2(D)$  a.s. ([78]). As we will be using the finite element method to approximate  $u^m$ , we know that the smoother  $u^m$  is, the faster its finite element approximation will converge. Furthermore, the FEM approximation of linear elliptic equations with  $L^2(D)$  right-hand side has been studied extensively and we can apply standard results.

## Elliptic regularity

The property  $u^m \in H^2(D)$  (a.s.) comes from a property of elliptic equations called elliptic regularity, which can be expressed in the following way (see the work by Hackbusch [80] or by Grisvard [78] for further details).

**Theorem 3.6** (elliptic regularity [78, 80, 126]). *Let  $s \geq -1$  be an integer and let  $Lu = f$  be a second order linear elliptic PDE in the form*

$$Lu = - \sum_{i,j=1}^d \frac{\partial}{\partial x_j} \left( a_{ij} \frac{\partial u}{\partial x_i} \right) + a_0 u = f, \quad a_0 \geq 0 \quad \text{a.e. in } D, \quad f \in H^s(D), \quad (3.32)$$

$$\sum_{i,j=1}^d a_{ij} \bar{\xi}_i \bar{\xi}_j \geq \alpha |\bar{\xi}|_2^2, \quad \alpha > 0, \quad \forall \mathbf{x} \in D, \quad \forall \bar{\xi} \in \mathbb{R}^d, \quad (3.33)$$

with homogeneous Dirichlet boundary conditions. If  $a_{ij} \in W^{s+1,\infty}(D)$ ,  $a_0 \in W^{s,\infty}(D)$  ( $a_0 \in L^\infty(D)$  if  $s = -1$ ) and the domain  $D$  is of class  $C^{s+1,1}$ , then  $u \in H^{s+2}(D) \cap H_0^1(D)$  and  $\|u\|_{H^{s+2}(D)} \leq c^*(s, d, D) \|f\|_{H^s(D)}$ . If the coefficients are constant<sup>2</sup> and the domain is of class  $C^{[s]+1,1}$  this result also holds for  $s$  real (for further details, see [80]).

If  $s < -1$  things get more complicated. However, it is still possible to obtain a similar result even in the case in which the right-hand side is a measure. In this case, the following result holds,

**Theorem 3.7** (elliptic regularity with measure data [38]). *Let  $d \in \{2, 3\}$  and consider the same assumptions as in Theorem 3.6 in the case in which  $s = -1$ , but take  $f$  to be a Radon measure instead (i.e.  $f \in (C_0^0(D))^*$ ), then a unique (very weak) solution  $u$  of the PDE exists and satisfies  $u \in W_0^{1,q}(D)$  with  $q = d/(d-1) - \epsilon/2$  for any  $\epsilon > 0$ , where by  $W^{1,q}(D)$  we indicate the usual Sobolev space of order  $q$  and*

$$\|u\|_{H^{2-d/2-\epsilon}(D)} \leq c^*(\epsilon, s, d, D) \|u\|_{W^{1,q}(D)} \leq \tilde{c}(\epsilon, s, d, D) \|f\|_{(C_0^0(D))^*}. \quad (3.34)$$

The constant  $\tilde{c}$  blows up as  $\epsilon \rightarrow 0$ .

*Proof.* The second inequality was proved by Casas in [38]. The first inequality is obtained through the use of the Sobolev imbedding theorem, see for Example [78].  $\square$

### 3.2.2 Main part of the proof

We can now finally proceed to the main part of the proof of Theorem 3.2. Recalling equation (3.9), the total error  $\mathbb{E}[\|u - u^h\|_{L^2(D)}^2]$ , can be written down as

$$\mathbb{E}[\|u - u_h^m\|_{L^2(D)}^2] \leq 2 \mathbb{E}[\|u - u^m\|_{L^2(D)}^2] + 2 \mathbb{E}[\|u^m - u_h^m\|_{L^2(D)}^2], \quad (3.35)$$

where we used the fact that  $u_h^m \equiv u_h$  by the definition of projected white noise (Definition 3.2). The proof of Theorem 3.2 is divided in two parts: in the first part we prove convergence of the first term on the RHS and in the second part we prove convergence of the second term.

---

<sup>2</sup>or under other suitable conditions, see [80].

### Convergence of $\mathbb{E}[||\mathbf{u} - \mathbf{u}^m||_{\mathbf{L}^2(\mathbf{D})}^2]$

**Lemma 3.8.** *Let  $V_h$  be the Lagrange finite element approximation subspace used to discretise (3.1). For any sufficiently small  $\epsilon > 0$ , if we let  $s = 2 - d/2 - \epsilon$ , we have for either  $p$  or  $h$  fixed,*

$$\mathbb{E}[||u - u^m||_{L^2(D)}^2]^{1/2} \leq \bar{c}(\epsilon, s, d, D)p^{-s}h^s, \quad \text{with } \bar{c}(\epsilon) \rightarrow \infty, \text{ as } \epsilon \rightarrow 0. \quad (3.36)$$

*Proof.* Let  $g(\mathbf{x}, \mathbf{y})$  be the Green's function of equation (3.5) over  $D$ . Then  $g(\mathbf{x}, \mathbf{y})$  satisfies, for any fixed  $\mathbf{y} \in D$ ,

$$g(\mathbf{x}, \mathbf{y}) - \kappa^{-2}\Delta_x g(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), \quad x \in D, \quad g(\mathbf{x}, \mathbf{y}) = 0, \quad \mathbf{x} \in \partial D, \quad (3.37)$$

where  $\delta$  is Dirac's delta distribution. By definition, the delta function is a Radon measure, i.e.  $\delta \in (C_0^0(D))^*$  (the continuity of test functions makes their point-evaluation well-defined). Theorem 3.7 then gives us  $g(\mathbf{x}, \mathbf{y}) \in H^s(D)$  for fixed  $\mathbf{y} \in D$ .

As  $\dot{\mathbf{W}}^m$  is in  $L^2(D)$  a.s. for finite  $m$ , then  $u^m \in H^2(D)$  a.s. for finite  $m$  and the  $L^2$  norm of  $u - u^m$  is finite a.s.. In particular, Fubini-Tonelli's theorem allows us to exchange the order of expectation and integration, giving us

$$\mathbb{E}[||u - u^m||_{L^2(D)}^2] = ||\mathbb{E}[|u - u^m|^2]||_{L^1(D)}. \quad (3.38)$$

By using the Green's function, we can write down analytic expressions for  $u$  and  $u^m$  (see also [32]),

$$u = \langle \dot{\mathbf{W}}, g(\mathbf{x}, \mathbf{y}) \rangle_x, \quad u^m = (\dot{\mathbf{W}}^m, g(\mathbf{x}, \mathbf{y}))_x = \langle \dot{\mathbf{W}}, P_h^x g(\mathbf{x}, \mathbf{y}) \rangle_x, \quad (3.39)$$

where by  $\langle \cdot, \cdot \rangle_x$ ,  $(\cdot, \cdot)_x$  and by  $P_h^x$  we mean that the pairing, inner product and orthogonal projection are taken with respect to the variable  $\mathbf{x}$ . This gives us, by linearity,

$$u - u^m = \langle \dot{\mathbf{W}}, g(\mathbf{x}, \mathbf{y}) - P_h^x g(\mathbf{x}, \mathbf{y}) \rangle_x = \langle \dot{\mathbf{W}} - \dot{\mathbf{W}}^m, g(\mathbf{x}, \mathbf{y}) \rangle_x. \quad (3.40)$$

Lemma 3.4 then gives us

$$\mathbb{E}[||u - u^m||_{L^2(D)}^2] = ||\mathbb{E}[|\langle \dot{\mathbf{W}} - \dot{\mathbf{W}}^m, g(\mathbf{x}, \mathbf{y}) \rangle_x|^2]||_{L^1(D)}, \quad (3.41)$$

$$\leq (c(s, d, D)p^{-s}h^s)^2 \int_D ||g(\mathbf{x}, \mathbf{y})||_{H^s, x}^2 d\mathbf{y}. \quad (3.42)$$

To conclude the proof, we need to show that the integral on the right-hand side is bounded. To do so, we will now show that  $||g(\mathbf{x}, \mathbf{y})||_{H^s, x}^2 \leq \tilde{c}(\epsilon)$  for all  $\mathbf{y} \in D$ , where  $\tilde{c}(\epsilon)$  does not depend on  $\mathbf{y}$ , but it blows up as  $\epsilon \rightarrow 0$ .

Combining theorems 3.6 and 3.7 we obtain

$$||g(\mathbf{x}, \mathbf{y})||_{H^s, x} \leq \tilde{c}(\epsilon, s, d, D)||\delta(\mathbf{x} - \mathbf{y})||_{(C_0^0(D))^*, x}, \quad (3.43)$$

where in the  $d = 1$  case we exploited the fact that the  $(C_0^0(D))^*$  norm is stronger than the  $H^{-1/2-\epsilon}(D)$  norm since  $H_0^{1/2+\epsilon}(D) \subset C_0^0(D) \subset (C_0^0(D))^* \subset H^{-1/2-\epsilon}(D)$ .

For all  $\mathbf{y} \in D$  we now have

$$\|\delta(\mathbf{x} - \mathbf{y})\|_{(C_0^0(D))^*, x} = \sup_{\|v\|_{C_0^0(D)} \leq 1} |\langle \delta(\mathbf{x} - \mathbf{y}), v(\mathbf{x}) \rangle_x| = \sup_{\|v\|_{C_0^0(D)} \leq 1} |v(\mathbf{y})| = 1. \quad (3.44)$$

Therefore,

$$\int_D \|g(\mathbf{x}, \mathbf{y})\|_{H^s(D), x}^2 d\mathbf{y} \leq \int_D (\tilde{c}(\epsilon, s, d, D) \cdot 1)^2 d\mathbf{y} = |D| \tilde{c}(\epsilon, s, d, D)^2, \quad (3.45)$$

and we conclude

$$\mathbb{E}[|u - u^m|^2_{L^2(D)}]^{1/2} \leq \bar{c}(\epsilon, s, d, D) p^{-s} h^s, \quad (3.46)$$

with  $\bar{c} = |D|^{1/2} c \tilde{c}$ , which is the proposition.  $\square$

*Remark 3.6.* Note that the convergence proof of the term  $\mathbb{E}[|u - u^m|^2_{L^2(D)}]$  ultimately boils down to the question ‘‘How regular is the Green’s function of the linear operator of the PDE and how accurately can one approximate it?’’. It is then easy to generalise this result to more general problems.

*Remark 3.7.* The fact that the constant appearing in the estimate of Lemma 3.8 blows up as  $\epsilon \rightarrow 0$  is the reason why the constant in the final estimate of Theorem 3.2 also blows up. This is an artefact of the technique used in this proof and can be removed. Bolin et al. in [27] solve this problem by introducing another ‘‘approximation’’ of white noise given by the expansion of white noise with respect to the discrete eigenbasis of the elliptic operator of (3.1) and showing that this is equivalent to  $W^m$  in  $L^2(\Omega, L^2(D))$ . The use of the discrete eigenbasis expansion allows one to derive estimates in which the  $\epsilon$  term does not appear, as also shown by Zhang et al. [182] for the continuous eigenbasis case.

### Convergence of $\mathbb{E}[|u^m - u_h^m|^2_{L^2(D)}]$

**Lemma 3.9.** *Let  $u_h^m$  be the finite element approximation to (3.5) obtained using Lagrange elements of degree  $p$  on each element and a quasi-uniform triangulation of  $D$  of maximum mesh size  $h$ . Then, for either fixed  $p$  or fixed  $h$ ,*

$$\mathbb{E}[|u^m - u_h^m|^2_{L^2(D)}]^{1/2} \leq \tilde{c}(d, D) p^{-(2-d/2)} h^{2-d/2}. \quad (3.47)$$

Note that  $m$  here is a function of  $h$  and  $p$  and  $m \rightarrow \infty$  and  $h/p \rightarrow 0$ .

*Proof.* We know from Lemma 3.5 that the projected white noise is in  $L^2(D)$  a.s. for finite  $m$ . As the domain is regular, elliptic regularity results (see Theorem 3.6) allow us to use the Aubin-Nitsche trick [12, 31] and obtain the convergence estimate

$$\|u^m - u_h^m\|_{L^2(D)} \leq C(d, D) p^{-2} h^2 \|u^m\|_{H^2(D)}, \quad \text{a.s.}, \quad (3.48)$$

for  $m$  large enough (i.e.  $h/p$  small enough). Again, elliptic regularity (see Theorem 3.6) gives us the bound

$$\|u^m\|_{H^2(D)} \leq c^*(d, D) \|\dot{W}^m\|_{L^2(D)}, \quad \text{a.s.} \quad (3.49)$$

Hence,

$$\|u^m - u_h^m\|_{L^2(D)} \leq c^*(d, D)C(d, D)p^{-2}h^2 \|\dot{W}^m\|_{L^2(D)}, \quad \text{a.s.} \quad (3.50)$$

We can then apply the expected value on both sides and use Lemma 3.5 to obtain

$$\mathbb{E}[\|u^m - u_h^m\|_{L^2(D)}^2]^{1/2} \leq c^*Cp^{-2}h^2 \mathbb{E}[\|\dot{W}^m\|_{L^2(D)}^2]^{1/2} = c^*Cp^{-2}h^2\sqrt{m}. \quad (3.51)$$

For a Lagrange finite element discretisation on a quasi-uniform mesh, the number  $m$  of basis functions needed is related to  $h$  and  $p$ , namely there exist  $C^* \in \mathbb{R}$  dependent only on  $D$  such that  $m \leq C^*p^d h^{-d}$ . This yields

$$\mathbb{E}[\|u^m - u_h^m\|_{L^2(D)}^2]^{1/2} \leq c^*Cp^{-2}h^2\sqrt{m} \leq \tilde{c}(d, D)p^{-(2-d/2)}h^{2-d/2}, \quad (3.52)$$

where  $\tilde{c} = c^*C\sqrt{C^*}$ , which is the proposition.  $\square$

*Remark 3.8.* Note that the problem of proving the order of convergence for this error term is strongly related to the question “How accurately can one solve the PDE with an  $L^2(D)$  forcing term by using finite elements?”.

### Conclusion of the proof

We can finally conclude the proof of Theorem 3.2.

*Proof of Theorem 3.2.* Let  $s = 2 - d/2 - \epsilon$ . The total error  $\mathbb{E}[\|u - u^h\|_{L^2(D)}^2]$ , can be written down as,

$$\mathbb{E}[\|u - u_h^m\|_{L^2(D)}^2] \leq 2\mathbb{E}[\|u - u^m\|_{L^2(D)}^2] + 2\mathbb{E}[\|u^m - u_h^m\|_{L^2(D)}^2] \quad (3.53)$$

$$\leq 2(\bar{c}(\epsilon, s, d, D)p^{-s}h^s)^2 + 2(\tilde{c}(d, D)p^{-(2-d/2)}h^{2-d/2})^2 \leq (\hat{C}(\epsilon, s, d, D)p^{-s}h^s)^2, \quad (3.54)$$

where  $\hat{C} = \sqrt{2}\max(\bar{c}, \tilde{c})$ . Taking the square root on both sides concludes the proof.  $\square$

## 3.3 Finite element convergence for problem (2.43): the case $k > 1$

The general case  $k > 1$  is just a slight generalisation of the  $k = 1$  case. In this section, we will prove the following result.

**Theorem 3.10.** *Let  $u_1^h \in V_h \subseteq H_0^1(D)$  be the FEM approximation of the solution of equation (3.1) obtained using continuous Lagrange elements of degree  $p$  on each cell and a quasi-uniform triangulation of  $D$  of maximum mesh size  $h$ . Let  $a(u, v) : H_0^1(D) \times H_0^1(D) \rightarrow \mathbb{R}$  be the bilinear form given by  $a(u, v) = (u, v) + \kappa^{-2}(\nabla u, \nabla v)$ . Consider the sequence of FEM approximations  $u_j^h \in V_h$  for  $j = 1, \dots, k$  such that*

$$a(u_{j+1}^h, v_h) = (u_j^h, v_h), \quad \text{a.s.} \quad \forall v_h \in V_h, \quad j = 1, \dots, k-1. \quad (3.55)$$

Let  $u_k$  be the solution of problem (2.43), and let  $s = 2k - d/2 - \epsilon$  for any  $\epsilon > 0$  sufficiently small, then  $u_k \in H^s(D)$  and for either fixed  $p$  or fixed  $h$ ,

$$\mathbb{E}[||u_k - u_k^h||_{L^2(D)}^2]^{1/2} \leq c(\epsilon, s, d, D)p^{-s}h^\mu, \quad \mu = \min(s, p+1), \quad (3.56)$$

where the constant  $c$  blows up as  $\epsilon \rightarrow 0$ .

The proof uses the same strategy as in the  $k = 1$  case: we split the problem in two by considering the FEM solution of problem (2.43) obtained by using projected white noise as the initial forcing term. We will now prove two lemmas, analogous to lemmas 3.8 and 3.9.

**Lemma 3.11.** *Let  $V_h$  be the Lagrange finite element approximation subspace used to approximate the solution of problem (2.43). Let  $u_k^m$  be the exact solution of problem (2.43) in which white noise is replaced with projected white noise. For any sufficiently small  $\epsilon > 0$ , if we let  $s = 2k - d/2 - \epsilon$  and  $\mu = \min(s, p+1)$ , we have, for either fixed  $p$  or fixed  $h$ ,*

$$\mathbb{E}[||u_k - u_k^m||_{L^2(D)}^2]^{1/2} \leq \bar{c}(\epsilon, s, d, D)p^{-s}h^\mu, \quad \text{with } \bar{c}(\epsilon) \rightarrow \infty, \text{ as } \epsilon \rightarrow 0. \quad (3.57)$$

*Proof.* The proof is substantially the same as the one of Lemma 3.8 and we therefore omit the details. The only difference is that if we let  $g_k$  be the Green's function of problem (2.43), then  $g_k$  is smoother than in Lemma 3.8 as here  $g_k \in H^s(D)$  (note that here  $s$  is larger) and Lemma 3.3 yields a higher convergence rate. The higher smoothness comes again from elliptic regularity (see theorems 3.6 and 3.7).  $\square$

**Lemma 3.12.** *Let  $V_h$  be the Lagrange finite element approximation subspace used to approximate the solution of problem (2.43). Let  $u_k^m$  be the exact solution of (2.43) in which white noise is replaced with projected white noise and let  $u_k^{m,h} \in V_h$  be its finite element approximation. Let  $\bar{s} = 2k - d/2$  and let  $\bar{\mu} = \min(\bar{s}, p+1)$ , then, for either fixed  $p$  or fixed  $h$ ,*

$$\mathbb{E}[||u_k^m - u_k^{m,h}||_{L^2(D)}^2]^{1/2} \leq \tilde{c}_k(d, D)p^{-\bar{s}}h^{\bar{\mu}}. \quad (3.58)$$

*Proof.* In the proof of Lemma 3.9 we showed that

$$||u_1^m||_{H^2(D)} \leq c_1^*(d, D)||\dot{W}^m||_{L^2(D)}, \quad \text{a.s.} \quad (3.59)$$

By repeatedly applying elliptic regularity estimates (see Theorem 3.6) it is then straightforward to show by induction that

$$||u_j^m||_{H^{2j}(D)} \leq c_j^*(d, D)||\dot{W}^m||_{L^2(D)}, \quad \text{a.s., } j = 1, \dots, k, \quad (3.60)$$

since  $u_{j-1}^m$  is the right-hand side for the second order elliptic PDE which is solved by  $u_j^m$ . We also know that the following standard FEM error bound holds for the FEM approximation of  $u_1^m$  (see [12, 31]),

$$||u_1^m - u_1^{m,h}||_{H^1(D)} \leq \bar{c}_1(d, D)p^{-1}h||u_1^m||_{H^2(D)} \quad \text{a.s..} \quad (3.61)$$

We now claim that for  $j = 1, \dots, k$ ,

$$\|u_j^m - u_j^{m,h}\|_{H^1(D)} \leq \bar{c}_j(d, D)p^{-(2j-1)}h^{\min(2j,p+1)-1}\|u_j^m\|_{H^{2j}(D)} \quad \text{a.s.} \quad (3.62)$$

Combining (3.60) and (3.62) for  $j = k$  we then obtain

$$\|u_k^m - u_k^{m,h}\|_{H^1(D)} \leq \bar{c}_k(d, D)c_k^*(d, D)p^{-(2k-1)}h^{\min(2k,p+1)-1}\|\dot{W}^m\|_{L^2(D)} \quad \text{a.s.} \quad (3.63)$$

As the domain is assumed to be regular, the Aubin-Nitsche argument (see [31, 152]) gives us

$$\|u_k^m - u_k^{m,h}\|_{L^2(D)} \leq \bar{c}_k(d, D)c_k^*(d, D)p^{-2k}h^{\min(2k,p+1)}\|\dot{W}^m\|_{L^2(D)} \quad \text{a.s.} \quad (3.64)$$

The proof is then concluded by using the same argument as in Lemma 3.9 by noting that the  $L^2$  norm of projected white noise is bounded by  $\sqrt{C^*}p^{d/2}h^{-d/2}$  a.s.. Squaring then both sides, taking the expectation and the square root gives the proposition.

The last thing to show is that the claim (3.62) actually holds. Consider equation (3.55): each approximate solution  $u_{j+1}^{m,h}$  is obtained by solving an elliptic PDE in which the right-hand side  $u_j^{m,h}$  is being approximated as well. The convergence of  $u_{j+1}^{m,h}$  will then also depend on how well is  $u_j^{m,h}$  approximated by the FEM scheme. In particular, Strang's first lemma (see [31]) gives us that

$$\|u_{j+1}^m - u_{j+1}^{m,h}\|_{H^1(D)} \leq \left(1 + \frac{c^a}{c_a}\right) \inf_{v_h \in V_h} \|u_{j+1}^m - v_h\|_{H^1(D)} + \|u_j^m - u_j^{m,h}\|_{H^{-1}(D)}, \quad (3.65)$$

a.s., where  $c^a = \max(1, \kappa^{-2})$  and  $c_a = \min(1, \kappa^{-2})$  are the continuity and coercivity constants of  $a(\cdot, \cdot)$  respectively. It is a standard result (see [12, 31, 152]) that the first term on the right is a.s. bounded by

$$\inf_{v_h \in V_h} \|u_{j+1}^m - v_h\|_{H^1(D)} \leq c_{1,j+1}(d, D)p^{-(2(j+1)-1)}h^{\min(2(j+1), p+1)-1}\|u_{j+1}^m\|_{H^{2(j+1)}(D)}. \quad (3.66)$$

For the second term, we use a duality argument. Let  $\phi \in H_0^1(D)$  and let  $w_\phi$  be the solution of the adjoint problem  $a(v, w_\phi) = (v, \phi)$  for each  $v \in H_0^1$ . Then, Galerkin orthogonality gives: for any  $w_h \in V_h$ ,

$$\begin{aligned} (u_j^m - u_j^{m,h}, \phi) &= a(u_j^m - u_j^{m,h}, w_\phi) = a(u_j^m - u_j^{m,h}, w_\phi - w_h) \\ &\leq c^a \|u_j^m - u_j^{m,h}\|_{H^1(D)} \|w_\phi - w_h\|_{H^1(D)}, \quad \text{a.s.} \end{aligned} \quad (3.67)$$

Standard convergence estimates and elliptic regularity [12, 31, 152] give us that

$$\|w_\phi - w_h\|_{H^1(D)} \leq c_3(d, D)p^{-2}h^2\|w_\phi\|_{H^3(D)} \leq c_4(d, D)p^{-2}h^2\|\phi\|_{H^1(D)} \quad \text{a.s.} \quad (3.68)$$

We now assume by an induction argument that the claim (3.62) holds for  $j$ . Combining this and equations (3.67) and (3.68) we obtain

$$(u_j^m - u_j^{m,h}, \phi) \leq c^a c_4(d, D)\bar{c}_j(d, D)p^{-(2(j+1)-1)}h^{\min(2(j+1), p+3)-1}\|\phi\|_{H^1(D)}, \quad \text{a.s.} \quad (3.69)$$

Taking the supremum over all  $\phi \in H_0^1(D)$  we get a bound for the  $H^{-1}(D)$  norm of  $u_j^m - u_j^{m,h}$ , namely,

$$\|u_j^m - u_j^{m,h}\|_{H^{-1}(D)} \leq c^a c_4(d, D) \bar{c}_j(d, D) p^{-(2(j+1)-1)} h^{\min(2(j+1), p+3)-1}, \text{ a.s.} \quad (3.70)$$

Using this with equation (3.65) and (3.66) we prove the case  $j + 1$  and we conclude the induction step. Since the base case of the induction argument is given by equation (3.61), claim (3.62) is proved.  $\square$

We now have all the ingredients to prove Theorem 3.10.

*Proof of Theorem 3.10.* The proof is the same as in Theorem 3.2: the proposition (3.56) comes from splitting the error  $\|u_k - u_k^{m,h}\|_{L^2(D)}$  in the two terms  $\|u_k - u_k^m\|_{L^2(D)}$  and  $\|u_k^m - u_k^{m,h}\|_{L^2(D)}$  and by noting that since projected white noise coincides with white noise if applied to any  $v_h \in V_h$  we have  $u_k^h \equiv u_k^{m,h}$ . The two error terms can be bounded with the convergence estimates in lemmas 3.11 and 3.12 respectively.

The result  $u_k \in H^s(D)$  comes by noting that  $u_1 \in H^{2-d/2-\epsilon}(D)$  (see Theorem 3.1) and by repeatedly applying elliptic regularity results to the sequence of problems generating  $u_2, \dots, u_k$  in problem (2.43).  $\square$

# Chapter 4

## Efficient white noise sampling and coupling for MLMC

The contents of this chapter are taken from our paper [43] and are novel unless otherwise indicated.

All the numerical results in this chapter have been produced with the software FEMLMC, a software for parallel forward uncertainty quantification for PDEs with random coefficients entirely written by the author that contains all the new methods presented in this thesis.

In Chapter 2 we introduced the SPDE approach to Matérn field sampling and in Chapter 3 we showed that the finite element method converges at the essentially optimal rate when applied to the white noise PDEs (3.1) and (2.43). In this chapter, we investigate the use of the SPDE approach within a non-nested MLMC framework and we present an efficient algorithm for the sampling and MLMC coupling of white noise realisations.

### 4.1 Background and existing literature

Let us recall the white noise PDE (2.41),

$$(\mathcal{I} - \kappa^{-2} \Delta)^k u(\mathbf{x}, \omega) = \eta \dot{\mathbf{W}}, \quad \mathbf{x} \in \mathbb{R}^d, \quad \omega \in \Omega, \quad \nu = 2k - d/2 > 0. \quad (4.1)$$

The main focus of this chapter is the generation of white noise samples  $\dot{\mathbf{W}}(\cdot, \omega)$  for a given sample point  $\omega \in \Omega$ . More precisely, we study the efficient sampling of the action  $\langle \dot{\mathbf{W}}, v_h \rangle(\omega)$  of white noise onto a FEM test function  $v_h$ . While solving equations (3.1) and (2.43) is relatively straightforward, the sampling of white noise realisations is not as it requires the sampling of a Gaussian vector with a finite element mass matrix  $M$  as covariance. If the finite element spaces involved are other than piecewise constants,  $M$  will be sparse, but not diagonal. Hence, its Cholesky factor is usually dense and the sampling requires an offline computational and memory storage cost of  $O(m^3)$  for the factorisation and an online cost of  $O(m^2)$  for each sample.

To resolve this challenge, different approaches have been adopted in the literature. Generally, the idea has been to use a diagonal mass matrix instead; i.e. an approximate

representation using piecewise constants or mass-lumping. Osborn et al. [132] use a two-field reformulation of (4.1) for  $k = 1$  with Raviart-Thomas elements combined with piecewise constants, while Lindgren et al. [116] use continuous Lagrange elements and mass lumping. Both methods compute (or approximate) the action of white noise on the FEM test functions. Another option, adopted in [51, 52, 141], is to approximate the white noise itself by a piecewise constant random function that converges in an appropriate weak sense to the exact white noise.

The sampling becomes more complicated when the Matérn field  $u$  is needed within a multilevel Monte Carlo framework [68, 69] which requires the coupling of the field between different approximation levels (i.e. the same sample point  $\omega$  must be used on both levels). In turn, this requires the white noise samples on each level to be coupled. Drzisga et al. [51] enforce this coupling in the nested grid case with the use of a piecewise constant approximation of white noise [149]. Osborn et al. [132] present a technique that enforces the coupling between nested meshes by using techniques from cell-based algebraic multigrid (AMG). Their approach does not require a user-provided hierarchy of nested grids as the hierarchy is constructed algebraically. This operation aggregates the cells of a single user-provided grid into clusters which then constitute the cells of the coarse meshes. The resulting aggregated meshes are non-simplicial. Furthermore, Osborn et al. [133] use a hierarchy of nested structured grids on which they enforce the white noise coupling and solve the SPDE (4.1). The techniques used for the coupling are the same as presented in [132]. The sampled Matérn fields are then transferred to a non-nested agglomerated mesh of the domain of interest via a Galerkin projection.

The main new contributions in this chapter are the following. First, we present a sampling technique for white noise that is exact and that is applicable for a wide range of finite element families, including all types of Lagrange elements. Our technique does not require the expensive factorisation of a global mass matrix or a costly two-field splitting of the Laplacian and has linear complexity in the number of degrees of freedom, hence justifying the complexity results in Table 2.1. Second, we introduce a technique for coupling white noise between nested or non-nested meshes, applicable for the same class of finite element families. If non-nested meshes are used, this coupling technique requires the use of a supermesh construction [59, 60, 61]. Third, we prove that the number of cells of a supermesh between two quasi-uniform meshes constructed with a local supermeshing algorithm [59] is bounded by the sum of the number of cells of the parent meshes, improving the complexity bounds of the algorithm. Finally, the existing literature generally focuses on white noise coupling in the  $h$ -refinement case, i.e. when the MLMC hierarchy is defined by meshes of decreasing cell size [39, 41]. In this chapter we also consider the case in which the MLMC levels are defined by increasing the polynomial degree of the FEM interpolant ( $p$ -refinement).

Although Osborn et al. [132] also work with non-nested meshes, our approach differs significantly from theirs. Osborn et al. start from one single mesh and algebraically coarsen it to obtain the grid hierarchy. The MLMC levels are thus generated algebraically. Our approach operates on a given arbitrary mesh hierarchy and the MLMC levels are defined geometrically. In our case, every mesh in the hierarchy is

simplicial, and it is thus possible to use standard FEM error estimates (if available) to estimate *a priori* the MLMC convergence parameters [39, 166].

We adopt the same embedded domain strategy as Osborn et al. [132]. The advantage of this strategy is that the sampled Matérn field can be transferred to the computational domain of interest exactly and at negligible cost. However, in practical applications defined over complex geometries, a sequence of nested meshes might not be available, making the white noise coupling challenging. This motivated us to design an algorithm that can be used to enforce the coupling between non-nested meshes as well.

The rest of the chapter is structured as follows. In Section 4.2 we describe the white noise sampling problem for the cases where both independent and coupled realisations are needed. In Section 4.4 we describe our new sampling technique that allows the sampling of independent and coupled white noise realisations efficiently. In Section 4.5 we present numerical results corroborating the theoretical results and demonstrating the performance of the technique. Finally we summarise the results of the chapter in Section 4.6.

## 4.2 The white noise sampling problem

In this section we describe the practical aspects of the numerical solution of (4.1) when either independent (standard Monte Carlo) or coupled (MLMC) Matérn field samples are needed. As we will see, the main complication lies in the sampling of white noise realisations.

### 4.2.1 Finite element solution of elliptic PDEs with white noise forcing

The solutions of the linear elliptic PDE (4.1) correspond to a Matérn field with covariance given by (2.27). As the main focus of this chapter is on white noise sampling, we will restrict our attention to the  $k = 1$  case and we will set  $\eta = 1$  from now on for simplicity, although our white noise sampling method generalises to any  $k > d/4$ .

Solving (4.1) over the whole of  $\mathbb{R}^d$  is generally not feasible. Instead,  $\mathbb{R}^d$  is typically truncated to a bounded domain  $D \subset\subset \mathbb{R}^d$  and some boundary conditions are chosen, usually homogeneous Neumann or Dirichlet [27, 116]. In what follows, we assume that the Matérn field sample is needed on a domain  $G \subset\subset D$  (recall Figure 3.1). If  $D$  is sufficiently large in the sense that the distance between  $\partial D$  and  $\partial G$  is larger than the correlation length  $\lambda$  then the error introduced by truncating  $\mathbb{R}^d$  to  $D$  is negligible (cf. Remark 2.12). After truncating the domain, (4.1) for  $k = 1$  reduces to

$$\begin{aligned} u - \kappa^{-2} \Delta u &= \dot{W} && \text{in } D, \\ u &= 0 && \text{on } \partial D. \end{aligned} \tag{4.2}$$

Existence and uniqueness of solutions to (4.2) was proven in [20] and in [32].

We will solve (4.2) using the finite element method, as described in Chapter 3. Let  $V_h = \text{span}(\phi_1, \dots, \phi_m) \subseteq H_0^1(D)$  be a continuous Lagrange approximation subspace. A discrete weak form of (4.2) then reads: find  $u_h \in V_h$  such that

$$(u_h, v_h) + \kappa^{-2}(\nabla u_h, \nabla v_h) = \langle \dot{W}, v_h \rangle \quad \text{for all } v_h \in V_h. \quad (4.3)$$

The coefficients of the basis function expansion for  $u_h$ , i.e. the  $u_i$  such that  $u_h = \sum_{i=1}^m u_i \phi_i$ , are given by the solution of a linear system

$$A\mathbf{u} = \mathbf{b}, \quad \text{with} \quad A_{ij} = (\phi_i, \phi_j) + \kappa^{-2}(\nabla \phi_i, \nabla \phi_j), \quad b_i = \langle \dot{W}, \phi_i \rangle. \quad (4.4)$$

This linear system (4.4) can be solved in  $O(m)$  time by using an optimal solver such as full multigrid.

*Remark 4.1.* In the general  $k > 1$ ,  $k \in \mathbb{N}$  case, then we must solve equation (2.43). In this case the elliptic operator is the same in all equations and the same finite element basis and solver can be reused to compute all the  $u_j$ .

By Definition 2.6,  $\mathbf{b}$  satisfies

$$\mathbf{b} \sim \mathcal{N}(0, M), \quad M_{ij} = (\phi_i, \phi_j), \quad (4.5)$$

i.e.  $\mathbf{b}$  is a zero-mean Gaussian vector with the finite element mass matrix  $M$  as covariance matrix. Sampling white noise realisations can thus be accomplished by sampling a Gaussian vector of mass matrix covariance.

In Section 4.4, we present a factorisation of  $M$  in the form  $HH^T$  (cf. (2.32)) that is both sparse and computationally efficient to compute, thus allowing for efficient sampling of white noise.

#### 4.2.2 Multilevel white noise sampling/white noise coupling condition

We now consider the case in which coupled Matérn field realisations are needed in a MLMC setting, i.e. we want to draw samples of  $u_\ell(\mathbf{x}, \omega)$  and  $u_{\ell-1}(\mathbf{x}, \omega)$  at two different levels of accuracy  $\ell$  and  $\ell - 1$  for the same  $\omega \in \Omega$ . To understand the challenge, we first recall the MLMC estimator 2.15: let  $\mathcal{P}(u) = P(\omega)$  be a functional of interest depending on the solution  $u$  of (4.2), then the MLMC estimator for  $\mathbb{E}[P]$  reads

$$\mathbb{E}[P] \approx \mathbb{E}[P_L] \approx \sum_{\ell=1}^L \left[ \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} (P_\ell(\omega_\ell^n) - P_{\ell-1}(\omega_\ell^n)) \right], \quad (4.6)$$

where  $P_\ell = \mathcal{P}(u_\ell)$  for all  $\ell = 1, \dots, L$ .

The increased efficiency of MLMC with respect to standard Monte Carlo relies on the assumption that on fine levels (large  $\ell$ ) the variance is small due to the fact that the levels are coupled, i.e. the sample point  $\omega_\ell^n$  is the same for both  $P_\ell(\omega_\ell^n)$  and  $P_{\ell-1}(\omega_\ell^n)$ . The coupling makes  $P_\ell(\omega_\ell^n)$  and  $P_{\ell-1}(\omega_\ell^n)$  strongly correlated. This diminishes the

variance of their difference, and therefore fewer samples are required to estimate the expected value. In fact, MLMC can be seen as a variance reduction technique in which the coupling between the levels is one of the key elements. If the coupling is not enforced correctly so that the samples of  $P_\ell$  and  $P_{\ell-1}$  become independent, then the variance of each term of the telescoping sum in (2.15) increases, significantly harming its efficiency and convergence properties.

Since the only stochastic element present in (4.2) is white noise, it is sufficient to use the same white noise sample on both levels to enforce the coupling requirement. More precisely, let  $V^\ell$  and  $V^{\ell-1}$  be the finite element spaces on level  $\ell$  and  $\ell-1$  respectively for  $\ell > 1$ . We consider the following two variational problems coupled by a common white noise sample: find  $u_\ell \in V^\ell = \text{span}(\phi_1^\ell, \dots, \phi_{m_\ell}^\ell)$  and  $u_{\ell-1} \in V^{\ell-1} = \text{span}(\phi_1^{\ell-1}, \dots, \phi_{m_{\ell-1}}^{\ell-1})$  such that for  $\omega_\ell^n \in \Omega$

$$(u_\ell, v_\ell) + \kappa^{-2}(\nabla u_\ell, \nabla v_\ell) = \langle \dot{W}, v_\ell \rangle (\omega_\ell^n), \quad \text{for all } v_\ell \in V^\ell, \quad (4.7)$$

$$(u_{\ell-1}, v_{\ell-1}) + \kappa^{-2}(\nabla u_{\ell-1}, \nabla v_{\ell-1}) = \langle \dot{W}, v_{\ell-1} \rangle (\omega_\ell^n), \quad \text{for all } v_{\ell-1} \in V^{\ell-1}. \quad (4.8)$$

where the terms on the right-hand side are coupled in the sense that they are centred Gaussian random variables with covariance  $\mathbb{E}[\langle \dot{W}, v_l \rangle \langle \dot{W}, v_s \rangle] = (v_l, v_s)$  for  $l, s \in \{\ell, \ell-1\}$ , as given by Definition 2.6.

Let  $\mathbf{u}_\ell \in \mathbb{R}^{m_\ell}$  and  $\mathbf{u}_{\ell-1} \in \mathbb{R}^{m_{\ell-1}}$  be the vectors of the finite element expansion coefficients of  $u_\ell$  and  $u_{\ell-1}$ , respectively. Following the same approach as in Section 4.2.1, we note that the coefficient vectors solve the following block-diagonal linear system,

$$\left[ \begin{array}{c|c} A^\ell & 0 \\ \hline 0 & A^{\ell-1} \end{array} \right] \left[ \begin{array}{c} \mathbf{u}_\ell \\ \mathbf{u}_{\ell-1} \end{array} \right] = \left[ \begin{array}{c} \mathbf{b}_\ell \\ \mathbf{b}_{\ell-1} \end{array} \right]. \quad (4.9)$$

Alternatively, by letting  $\mathbf{u} = [\mathbf{u}_\ell, \mathbf{u}_{\ell-1}]^T$ ,  $\mathbf{b} = [\mathbf{b}_\ell, \mathbf{b}_{\ell-1}]^T$  and  $A = \text{diag}(A^\ell, A^{\ell-1})$ , we can write this as

$$A\mathbf{u} = \mathbf{b}. \quad (4.10)$$

This system can be solved in linear time with an optimal solver [54].

Furthermore, by Definition 2.6,

$$\mathbf{b} \sim \mathcal{N}(0, M), \quad (4.11)$$

where  $M$  can be expressed in block structure as

$$M = \left[ \begin{array}{c|c} M^\ell & M^{\ell, \ell-1} \\ \hline (M^{\ell, \ell-1})^T & M^{\ell-1} \end{array} \right], \quad \text{with} \quad M_{ij}^{\ell, k} = (\phi_i^\ell, \phi_j^k) \text{ and } M_{ij}^k = (\phi_i^k, \phi_j^k). \quad (4.12)$$

If we were using independent white noise samples for (4.7) and (4.8), then the off-diagonal blocks of  $M$  would vanish. Conversely, the presence of the mixed mass matrix  $M^{\ell, \ell-1}$  stems from the use of the same white noise sample on both levels. For this reason, we will refer to equations (4.11) and (4.12) as the *coupling condition*.

Thus, the problem of sampling coupled Matérn fields in the context of MLMC again reduces to the sampling of a Gaussian vector with a mass matrix as covariance. However, two additional complications arise. First,  $M$  is potentially much larger and not necessarily of full rank (consider the case in which  $V^\ell = V^{\ell-1}$ , then  $M^\ell = M^{\ell-1} = M^{\ell,\ell-1}$ ). Second, to assemble  $M^{\ell,\ell-1}$  we need to compute integrals involving basis functions possibly defined over different, non-nested meshes, which is non-trivial. In Section 4.4, we present a sampling technique that addresses both issues. A supermesh construction [60, 61] is required in the non-nested mesh case.

#### 4.2.3 Embedded meshes and non-nested grids

We adopt the same embedded mesh strategy as presented by Osborn et al. [132]. We assume that the Matérn field sample is needed on a user-provided mesh  $G_h$  of the domain  $G$  and we take  $D$  to be a larger  $d$ -dimensional box such that the distance between  $\partial G$  and  $\partial D$  is at least  $\lambda$ . With modern meshing software, such as Gmsh [65], it is possible to then triangulate  $D$  and obtain a mesh  $D_h$  in such a way that  $G_h$  is nested within  $D_h$ , i.e., each cell and vertex of  $G_h$  is also a cell or vertex of  $D_h$ . We then refer to  $G_h$  as embedded in  $D_h$  or to  $G_h$  as an embedded mesh (in  $D_h$ ). The main advantage of an embedded  $G_h$  in  $D_h$  is that once (4.2) is solved on  $D_h$  the sampled Matérn field  $u$  can be exactly transferred onto  $G_h$  at negligible cost. Conversely, if  $G_h$  is not embedded in  $D_h$ , an additional interpolation step would be required, thus increasing the cost of each sample.

In the MLMC framework with  $h$ -refinement, we assume that we are given a possibly non-nested user-provided mesh hierarchy  $\{G_h^\ell\}_{\ell=1}^L$ . We accordingly generate a hierarchy of meshes  $\{D_h^\ell\}_{\ell=1}^L$  on which to perform the sampling. If the  $\{D_h^\ell\}_{\ell=1}^L$  are nested, then the techniques used in [51] and [132] can be used to couple the white noise between MLMC levels. However, in the case in which the user-provided meshes  $\{G_h^\ell\}_{\ell=1}^L$  are non-nested, these methods are not compatible with the embedded mesh strategy.

Clearly, non-nested grid hierarchies appear naturally in practical computations on complex geometries. For instance, grid hierarchies generated from CAD geometries or through coarsening of a single fine mesh are generally non-nested. Thus tackling couplings across non-nested meshes is crucial for non-trivial applications. As previously said, such couplings can be achieved at a small offline cost via a supermesh construction [60, 61].

### 4.3 Supermesh constructions between quasi-uniform meshes

In what follows we consider the general setting in which the MLMC levels are defined using  $h$ -refinement and the mesh hierarchy  $\{D_h^\ell\}_{\ell=1}^L$  is non-nested. At the end of the next section, we will provide some remarks on the simpler cases in which the function spaces that define the hierarchy are nested (e.g. the grids are nested or  $p$ -refinement is used).

Consider the case of sampling  $\mathbf{b} \sim \mathcal{N}(0, M)$ , where  $M \in \mathbb{R}^{m \times m}$  is given by (4.12). The assembly of the off-diagonal blocks of  $M$  requires the computation of inner products between basis functions of different FEM approximation subspaces. To address this problem, we use a *supermesh* construction defined as follows.

**Definition 4.1** (Supermesh, [60, 61]). Let  $D \subset \subset \mathbb{R}^d$  be an open domain and let  $\mathcal{T}_a$ ,  $\mathcal{T}_b$  be two tessellations of  $D$ . A supermesh  $S_h$  of  $\mathcal{T}_a$  and  $\mathcal{T}_b$  is a common refinement of  $\mathcal{T}_a$  and  $\mathcal{T}_b$ . More specifically,  $S$  is a triangulation of  $D$  such that:

1.  $\text{vertices}(\mathcal{T}_a) \cup \text{vertices}(\mathcal{T}_b) \subseteq \text{vertices}(S_h)$ ,
2.  $\text{volume}(e_S \cap e) \in \{0, \text{volume}(e_S)\}$  for all cells  $e_S \in S_h$ ,  $e \in (\mathcal{T}_a \cup \mathcal{T}_b)$ .

The first condition means that every parent mesh vertex must also be a vertex of the supermesh, while the second states that every supermesh cell is completely contained within exactly one cell of either parent mesh [61]. The supermesh construction is not unique [61]. We show an example of supermesh construction in Figure 4.1. Efficient algorithms for computing the supermesh are available [121].

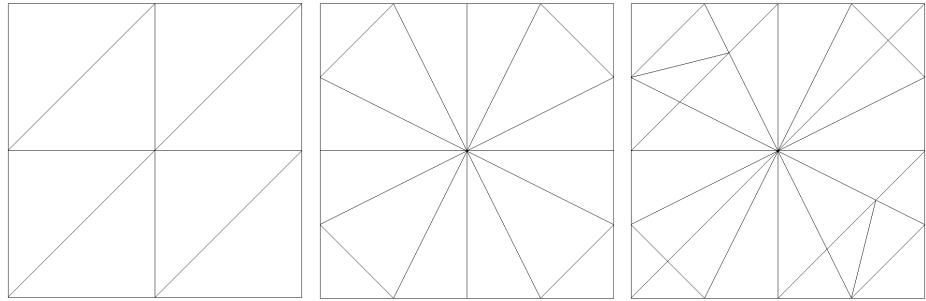


Figure 4.1: An example of a supermesh construction. The first two meshes on the left are the parent meshes and the mesh on the right is a supermesh. As stated in [61, Lemma 2], every supermesh cell is completely contained within a unique pair of parent mesh cells.

Evaluating (4.12) involves  $L^2$ -inner products of functions that are only *piecewise* polynomial on each cell of  $D_h^\ell$  and  $D_h^{\ell-1}$ . This lack of smoothness affects the convergence of standard quadrature schemes. The supermesh construction provides a resolution to this problem: on each cell of a supermesh of  $D_h^\ell$  and  $D_h^{\ell-1}$  the integrands are polynomial and standard quadrature schemes apply.

Overall, a key ingredient of our strategy for evaluating (4.12) is therefore to construct a supermesh of each pair of meshes  $D_h^\ell$ ,  $D_h^{\ell-1}$ . If the supermesh construction is performed with a local supermeshing algorithm, then its complexity is  $O(n_\ell + K)$ , where  $n_\ell$  is the number of cells of  $D_h^\ell$  and  $K = \mathcal{I}(D_h^\ell, D_h^{\ell-1})$  is the number of intersecting cells between  $D_h^\ell$  and  $D_h^{\ell-1}$  [59]. As stated in [61, Lemma 2], supermesh cells always lie within the intersection of a single pair of parent mesh cells and therefore the number of supermesh cells is proportional to the number of intersecting cells and is thus  $O(K)$  [61].

To our knowledge, the only bound for  $K$  available in the literature is given by  $K \leq c(d)n_\ell n_{\ell-1}$ , where  $n_{\ell-1}$  is the number of cells of  $D_h^{\ell-1}$  and  $c(d)$  is the minimum worst-case number of simplices that the intersection between two parent mesh cells can be triangulated into. For simplicial meshes we have  $c = 4$  in 2D and  $c = 45$  in 3D [59, 61]. In practice, this is a pessimistic bound and we will now show that under the additional assumption that the input meshes are quasi-uniform, the number of cells of the supermesh is bounded by a constant times the sum of the cells of the parent meshes,  $n_\ell + n_{\ell-1}$ . This in turn means that for practical computations the supermesh construction and the number of supermesh cells is  $O(n_\ell)$  (since here  $n_\ell > n_{\ell-1}$ ).

Let us first recall the definition of quasi-uniformity.

**Definition 4.2** (definition 4.4.13 in [31]). Let  $D$  be a given domain and let  $\{\mathcal{T}_{\hat{h}}\}$ , be a family of tessellations of  $D$  such that for  $0 < \hat{h} \leq 1$ ,

$$\max\{\text{diam } e : e \in \mathcal{T}_{\hat{h}}\} \leq \hat{h} \text{ diam } D, \quad (4.13)$$

where  $\text{diam } D$  is the diameter of  $D$ . The family is said to be quasi-uniform if there exists  $\infty > \hat{\rho} > 0$  (independent from  $\hat{h}$ ) such that

$$\min\{\text{diam } \underline{\mathcal{B}}_e : e \in \mathcal{T}_{\hat{h}}\} \geq \hat{\rho} \hat{h} \text{ diam } D, \quad (4.14)$$

where  $\underline{\mathcal{B}}_e$  is the largest ball contained in  $e$  such that  $e$  is star-shaped with respect to  $\underline{\mathcal{B}}_e$  (cf. definition 4.2.2 in [31]).

To simplify the exposition of what follows, it is more convenient to use the following property of quasi-uniform tessellations:

**Lemma 4.1.** *Let  $D$  be a given domain and let  $\{\mathcal{T}_h\}$  be a quasi-uniform family of tessellations. Then there exist  $h$  and  $\rho \in (0, \infty)$  with  $0 < h \leq c_d = \sqrt{2d/(d+1)}$  such that*

$$\max\{\text{diam } \bar{\mathcal{B}}_e : e \in \mathcal{T}_h\} \leq h \text{ diam } D, \quad (4.15)$$

where  $\bar{\mathcal{B}}_e$  is the smallest ball containing  $e$  and

$$\min\{\text{diam } \underline{\mathcal{B}}_e : e \in \mathcal{T}_h\} \geq \rho h \text{ diam } D, \quad (4.16)$$

where  $\underline{\mathcal{B}}_e$  is the largest ball contained in  $e$  such that  $e$  is star-shaped with respect to  $\underline{\mathcal{B}}_e$ .

*Proof.* Jung's Theorem [99, 100] states that for any compact set  $e \subset \mathbb{R}^d$ ,

$$\text{diam } \bar{\mathcal{B}}_e \leq \sqrt{\frac{2d}{d+1}} \text{ diam } e = c_d \text{ diam } e. \quad (4.17)$$

Therefore

$$\max\{\text{diam } \bar{\mathcal{B}}_e : e \in \mathcal{T}_h\} \leq c_d \max\{\text{diam } e : e \in \mathcal{T}_h\} \leq c_d \hat{h} \text{ diam } D, \quad (4.18)$$

where we have used equation (4.13) in the last step. Equation (4.14) also gives us,

$$\min\{\text{diam } \underline{\mathcal{B}}_e : e \in \mathcal{T}_h\} \geq \hat{\rho} \hat{h} \text{ diam } D. \quad (4.19)$$

Equations (4.18) and (4.19) are the same as (4.15) and (4.16) respectively after setting  $h = c_d \hat{h}$  and  $\rho = \hat{\rho}/c_d$ .  $\square$

In what follows we also need the following auxiliary lemma, stating that the constants  $h$  and  $\rho$  appearing in Lemma 4.1 also provide a lower and upper bound for the number of cells of a quasi-uniform mesh.

**Lemma 4.2.** *Let  $D_h$  be a quasi-uniform tessellation of a domain  $D \subset \mathbb{R}^d$  with  $n$  cells and let  $c_D = 2^{-d}|D|/(c_\pi(d)\text{diam}(D)^d)$  with  $c_\pi = 2$  in 1D,  $c_\pi = \pi$  in 2D and  $c_\pi = 4\pi/3$  in 3D, then*

$$c_D h^{-d} \leq n \leq c_D \rho^{-d} h^{-d}, \quad (4.20)$$

where  $h$  and  $\rho$  are the constants appearing in Lemma 4.1.

*Proof.* Let  $e_i \in D_h$  for  $i = 1, \dots, n$  be the cells of  $D_h$ . We compute a lower bound for  $n$  by noting that the measure of each cell is smaller or equal to the volume of the smallest ball containing it, which gives

$$|D| = \sum_i |e_i| \leq c_\pi 2^{-d} \sum_i \text{diam}(\bar{\mathcal{B}}_{e_i})^d \leq c_\pi 2^{-d} nh^d \text{diam}(D)^d, \quad (4.21)$$

where we have used equation (4.15) in the last step. The lower bound is obtained by solving for  $n$ . Similarly we obtain an upper bound by noting that the volume of each cell is larger than the volume of any ball it contains. This gives,

$$|D| = \sum_i |e_i| \geq c_\pi 2^{-d} \sum_i \text{diam}(\underline{\mathcal{B}}_{e_i})^d \geq c_\pi 2^{-d} n \rho^d h^d \text{diam}(D)^d, \quad (4.22)$$

where we used equation (4.16) in the last step. Solving for  $n$  yields an upper bound and concludes the proof.  $\square$

We now prove the following theorem, stating that the number of intersections  $K = O(n_\ell + n_{\ell-1})$ . To the author's knowledge, this result is novel.

**Theorem 4.3.** *Let  $A_h$  and  $B_h$  be two quasi-uniform tessellations of the same domain  $D$  with  $n_A$  and  $n_B$  be the number of cells of  $A_h$  and  $B_h$  respectively and let  $\mathcal{I}(A_h, B_h)$  be the number of intersecting pairs between  $A_h$  and  $B_h$ . There exists  $c(d, D) > 0$  independent from  $n_A$  and  $n_B$  such that  $\mathcal{I}(A_h, B_h) \leq c(d, D)(n_A + n_B)$ .*

*Proof.* Since  $A_h$  and  $B_h$  are quasi-uniform, we have by Lemma 4.1,

$$\max\{\text{diam } \bar{\mathcal{B}}_e : e \in A_h\} \leq h_A \text{diam } D = \tilde{h}_A, \quad (4.23)$$

$$\max\{\text{diam } \bar{\mathcal{B}}_e : e \in B_h\} \leq h_B \text{diam } D = \tilde{h}_B, \quad (4.24)$$

$$\min\{\text{diam } \underline{\mathcal{B}}_e : e \in A_h\} \geq \rho_A h_A \text{diam } D = \rho_A \tilde{h}_A, \quad (4.25)$$

$$\min\{\text{diam } \underline{\mathcal{B}}_e : e \in B_h\} \geq \rho_B h_B \text{diam } D = \rho_B \tilde{h}_B, \quad (4.26)$$

for some  $h_A, h_B \in (0, c_d]$ ,  $\rho_A, \rho_B > 0$  independent from  $h_A, h_B$ . For a given cell  $e_i \in A_h$ , let  $\mathcal{I}(B_h, e_i)$  be the number of cells of  $B_h$  that intersect with  $e_i$ , we then have that

$$\mathcal{I}(B_h, A_h) \leq n_A \max_i \mathcal{I}(B_h, e_i). \quad (4.27)$$

Now,  $\max_i \mathcal{I}(B_h, e_i)$  is bounded above by the maximum number  $\mathcal{P}(B_h, \mathcal{B}_{\tilde{d}})$  of cells of  $B_h$  that can be packed within a ball  $\mathcal{B}_{\tilde{d}}$  of diameter  $\tilde{d} = \tilde{h}_A + 2\tilde{h}_B$  without overlapping, since all cells  $e_i^A \in A_h$  can be entirely contained within a ball of diameter  $\tilde{h}_A$  and no cell  $e_j^B \in B_h$  can intersect  $e_i^A$  if the maximum distance between  $e_i^A$  and the points of  $e_j^B$  is larger than  $\tilde{h}_B$ , i.e.

$$|e_i^A \cap e_j^B| = 0, \quad \text{if } \max_{\mathbf{x} \in e_j^B} \text{dist}(\mathbf{x}, e_i^A) > \tilde{h}_B, \quad (4.28)$$

which gives,

$$|e_i^A \cap e_j^B| = 0, \quad \text{if } \exists \mathcal{B}_{\tilde{h}_B} \supseteq e_i^A \text{ s.t. } \max_{\mathbf{x} \in e_j^B} \text{dist}(\mathbf{x}, \mathcal{B}_{\tilde{h}_A}) > \tilde{h}_B. \quad (4.29)$$

From these considerations and (4.27) we get  $\mathcal{I}(B_h, A_h) \leq n_A \mathcal{P}(B_h, \mathcal{B}_{\tilde{d}})$ , where we can bound

$$\mathcal{P}(B_h, \mathcal{B}_{\tilde{d}}) \leq \mathcal{P}(\mathcal{B}_{\rho_B \tilde{h}_B}, \mathcal{B}_{\tilde{d}}), \quad (4.30)$$

where with abuse of notation we indicate with  $\mathcal{P}(\mathcal{B}_{\rho_B \tilde{h}_B}, \mathcal{B}_{\tilde{d}})$  the number of balls of diameter  $\rho_B \tilde{h}_B$  that can be packed within  $\mathcal{B}_{\tilde{d}}$  without overlapping. This bound holds since all cells of  $B_h$  entirely contain  $\mathcal{B}_{\rho_B \tilde{h}_B}$  by quasi-uniformity and Lemma 4.1. Note that  $\mathcal{B}_{\tilde{d}}$  is always larger than  $\mathcal{B}_{\rho_B \tilde{h}_B}$ . Finding the sharpest possible upper bound for  $\mathcal{P}(\mathcal{B}_{\rho_B \tilde{h}_B}, \mathcal{B}_{\tilde{d}})$  is a classical, yet extremely complicated problem in geometry called the ball packing problem (see e.g. [168] for a survey). A crude upper bound is given by

$$\mathcal{P}(\mathcal{B}_{\rho_B \tilde{h}_B}, \mathcal{B}_{\tilde{d}}) \leq \delta_d \frac{|\mathcal{B}_{\tilde{d}}|}{|\mathcal{B}_{\rho_B \tilde{h}_B}|} \leq \delta_d \left( \frac{\tilde{h}_A + 2\tilde{h}_B}{\rho_B \tilde{h}_B} \right)^d \leq \delta_d \rho_B^{-d} \left( 2 + \frac{h_A}{h_B} \right)^d \quad (4.31)$$

where  $\delta_d \leq 1$  is the packing density of congruent balls in  $\mathbb{R}^d$  and we removed the tildes since the diam  $D$  term simplifies out in the ratio. Putting everything together yields

$$\mathcal{I}(B_h, A_h) \leq \delta_d \rho_B^{-d} \left( 2 + \frac{h_A}{h_B} \right)^d n_A. \quad (4.32)$$

By Lemma 4.2, we have that

$$c_D h_A^{-d} \leq n_A, \quad \text{and} \quad n_B \leq c_D \rho_B^{-d} h_B^{-d}. \quad (4.33)$$

We can now use the above to compute an upper bound for  $h_A/h_B$  as follows. First compute an upper bound for  $h_A$  and a lower bound for  $h_B$ ,

$$h_A \leq \left( \frac{n_A}{c_D} \right)^{-1/d} \quad \text{by (4.33) left.} \quad h_B \geq \frac{1}{\rho_B} \left( \frac{n_B}{c_D} \right)^{-1/d} \quad \text{by (4.33) right.} \quad (4.34)$$

Then combine the above into

$$\frac{h_A}{h_B} = \frac{1}{h_B} h_A \leq \rho_B \left( \frac{n_B}{c_D} \right)^{1/d} \left( \frac{n_A}{c_D} \right)^{-1/d} = \rho_B \left( \frac{n_B}{n_A} \right)^{1/d}. \quad (4.35)$$

Plugging this into (4.32) we obtain

$$\mathcal{I}(B_h, A_h) \leq \delta_d \rho_B^{-d} \left( 2n_A^{1/d} + \rho_B n_B^{1/d} \right)^d \leq c(d, D)(n_A + n_B), \quad (4.36)$$

which concludes the proof. Here  $c(d, D) = \delta_d \bar{C}$  for some constant  $\bar{C} > 0$  such that

$$\left( 2\rho_B^{-1} n_A^{1/d} + n_B^{1/d} \right)^d \leq \bar{C}(n_A + n_B). \quad (4.37)$$

□

**Corollary 4.4.** *Let  $A_h$  and  $B_h$  be two quasi-uniform tessellations of the same domain  $D$  with  $n_A$  and  $n_B$  be the number of cells of  $A_h$  and  $B_h$  respectively and let  $S_h$  be a supermesh with  $n$  cells constructed from  $A_h$  and  $B_h$  via a local supermeshing algorithm [59]. There exists  $C(d, D) > 0$  independent from  $n_A$  and  $n_B$  such that  $n \leq C(d, D)(n_A + n_B)$ .*

*Proof.* For a local supermeshing algorithm (cf. [59]) we have  $n \leq c(d) \mathcal{I}(B_h, A_h)$ , where  $c(d) > 0$  is the minimum worst-case number of simplices that the intersection between two cells of  $A_h$  and  $B_h$  can be triangulated into [61]. The proposition follows from Theorem 4.3. □

*Remark 4.2.* The constant in the bound derived in Theorem 4.3 is not sharp. However, a sharper constant can be estimated in practice, cf. Section 4.5.

*Remark 4.3.* Generalising the theorem to non-nested meshes of different overlapping domains is straightforward as it is sufficient to bound the number of intersections in the region in which the meshes are overlapping.

We therefore conclude that the supermeshing cost and the number of cells  $n$  of a supermesh constructed from  $D_h^\ell$  and  $D_h^{\ell-1}$  by using a local supermeshing algorithm [59] are both  $O(n_\ell)$  since  $n_\ell > n_{\ell-1}$ . This is an important result because the sampling strategy for coupled white noise realisations that we present in the next section has cost complexity linear in the number of supermesh cells. Consequently, the method we propose is linear in  $n_\ell$ .

## 4.4 White noise sampling

In this section we introduce a new technique for sampling white noise efficiently. We first address the basic case in which independent white noise samples are needed before considering the more complicated case in which coupled samples are required.

#### 4.4.1 Sampling of independent white noise realisations

As discussed in the previous section, the sampling of independent white noise realisations defined over a meshed domain can be cast as the sampling of a Gaussian vector  $\mathbf{b}$  of covariance matrix given by a finite element mass matrix  $M \in \mathbb{R}^{m \times m}$ . In turn, efficient sampling of such a Gaussian vector typically involves computing a factorisation of  $M = HH^T$ . If a Cholesky factorisation is used, such sampling may become costly, with a  $O(m^3)$  factorisation cost and  $O(m^2)$  cost per sample. In what follows, we present an alternative factorisation strategy which has  $O(m)$  fixed cost and  $O(m)$  cost per sample.

The core idea is to work cell-wise instead of factorizing a global mass matrix. To illustrate, consider a standard finite element assembly of  $M$  over a mesh with  $n$  cells and  $m_e$  degrees of freedom on each cell. Local mass matrices  $M_e$  of size  $m_e \times m_e$  are computed on each mesh cell  $e$  before aggregation to form the global mass matrix  $M$ . The overall assembly operation can be written in matrix form,

$$M = L^T \text{diag}_e(M_e) L, \quad (4.38)$$

(see e.g. [173]), where  $\text{diag}_e(M_e)$  is a block diagonal matrix of size  $nm_e \times nm_e$  with the local mass matrices on the diagonal and  $L$  is a Boolean assembling matrix of size  $nm_e \times m$  such that  $L^T = [L_1^T \dots L_n^T]$  and the  $L_e$  are Boolean matrices of size  $m_e \times m$  that encode the local-to-global map. Note that each row of  $L$  has exactly one non-zero entry [173].

We can now factorise each local mass matrix  $M_e$  independently with a standard Cholesky factorisation to obtain  $M_e = H_e H_e^T$  for each cell  $e$ . We then have

$$M = L^T \text{diag}_e(H_e H_e^T) L = (L^T \text{diag}_e(H_e))(L^T \text{diag}_e(H_e))^T = HH^T, \quad (4.39)$$

with  $H \equiv L^T \text{diag}_e(H_e)$ , and we can sample  $\mathbf{b}$  by computing

$$\mathbf{b} = H\mathbf{z}, \quad \text{with } \mathbf{z} \sim \mathcal{N}(0, I), \quad \mathbf{z} \in \mathbb{R}^{m_e n}, \quad (4.40)$$

since, cf. (2.32),

$$\mathbb{E}[\mathbf{b}\mathbf{b}^T] = H \mathbb{E}[\mathbf{z}\mathbf{z}^T] H^T = (L^T \text{diag}_e(H_e)) I (L^T \text{diag}_e(H_e))^T = M. \quad (4.41)$$

*Remark 4.4.* This sampling strategy allows the splitting of a large global sampling problem into separate small local sampling problems. In fact, if for each cell  $e$  we let  $\mathbf{z}_e \sim \mathcal{N}(0, I)$  be a small standard Gaussian vector of length  $m_e$ , we can rewrite  $\mathbf{b} = H\mathbf{z}$  as

$$\mathbf{b} = H\mathbf{z} = \sum_{e=1}^n L_e^T H_e \mathbf{z}_e = \sum_{e=1}^n L_e^T \mathbf{b}_e, \quad (4.42)$$

where  $\mathbf{b}_e \sim \mathcal{N}(0, M_e)$  is sampled locally. The problem of sampling a global mass matrix covariance Gaussian vector then eventually reduces to the sampling of  $n$  independent local mass matrix covariance Gaussian vectors. This sampling approach is therefore trivially parallelisable.

Note that this sampling strategy is efficient since the local Cholesky factorisations can be computed in  $O(m_e^3 n)$  time and the  $L_e$  factors can be applied matrix-free for a total  $O(m_e^3 n)$  factorisation cost and an  $O(m_e^2 n)$  memory and sampling cost.

*Remark 4.5.* In the case in which the transformation to the reference cell is affine (such as with Lagrange elements on simplices) this operation can be made much more efficient by noting that the local mass matrices on each cell are always the same up to a multiplicative factor, namely  $M_e/|e| = \text{const}$  for all  $e$ , where  $|e|$  is the measure of the cell. It is therefore sufficient to factorise a single local mass matrix and to store its Cholesky factor, yielding a negligible  $O(m_e^3)$  and  $O(m_e^2)$  factorisation and memory cost respectively.

We note that the standard Gaussian vector  $\mathbf{z}$  used to compute  $\mathbf{b}$  is of size  $m_e n$  which is larger than if a Cholesky factorisation was used (a Cholesky factor would be of size  $m \times m$ , yielding a standard Gaussian vector of length  $m$ ). In fact, unlike the Cholesky factor, the matrix  $H$  here is not square. However, in comparison to the cost of solving (4.2), the sampling cost of the extra Gaussian variables is negligible.

#### 4.4.2 Sampling coupled white noise realisations for MLMC

We now turn to consider the case of sampling coupled white noise. In what follows we consider the general setting in which the MLMC levels are defined using  $h$ -refinement and the mesh hierarchy is non-nested. At the end of the section, we provide some remarks on the simpler cases in which the function spaces that define the hierarchy are nested (e.g. the grids are nested or  $p$ -refinement is used).

##### Supermesh construction and global mass matrix assembly

We now consider the case of sampling  $\mathbf{b} \sim \mathcal{N}(0, M)$ , where  $M \in \mathbb{R}^{m \times m}$  is given by (4.12). The assembly of the off-diagonal blocks of  $M$  requires the computation of inner products between basis functions of different FEM approximation subspaces. As stated in the previous section we use a supermesh construction between each pair of meshes  $D_h^\ell, D_h^{\ell-1}$  to address this problem.

Note that each supermesh cell lies in the intersection of exactly one pair  $(e_\ell, e_{\ell-1})$  of parent mesh cells  $e_\ell \in D_h^\ell, e_{\ell-1} \in D_h^{\ell-1}$ , hence we only need to account for the basis functions that are non-zero over  $e_\ell$  and  $e_{\ell-1}$ . Let  $m_{e_\ell}$  and  $m_{e_{\ell-1}}$  denote the number of degrees of freedom defined by the finite element spaces  $V^\ell$  and  $V^{\ell-1}$  over cells  $e_\ell$  and  $e_{\ell-1}$  respectively. Then, only the inner products between  $m_e = m_{e_\ell} + m_{e_{\ell-1}}$  basis functions will be non-zero.

We can thus assemble  $M$  given by (4.12) by the following two-step algorithm.

1. Let  $n$  be the number of supermesh cells. For each supermesh cell  $e$ , use quadrature rules over  $e$  to compute the local mass matrix

$$M_e = \left[ \begin{array}{c|c} M_e^\ell & M_e^{\ell, \ell-1} \\ \hline (M_e^{\ell, \ell-1})^T & M_e^{\ell-1} \end{array} \right], \quad (M_e^{\ell, \ell-1})_{ij} = \int_e \phi_i^\ell \phi_j^{\ell-1} \, dx, \quad (4.43)$$

where  $\{\phi_i^\ell\}_{i=1}^{m_{e_\ell}}$  and  $\{\phi_j^{\ell-1}\}_{j=1}^{m_{e_{\ell-1}}}$  are sets of the basis functions of  $V^\ell$  and  $V^{\ell-1}$  respectively that have non-zero support over  $e$ .  $M_e$  is of size  $m_e \times m_e$ ,  $M_e^\ell$  is of size  $m_{e_\ell} \times m_{e_\ell}$ ,  $M_e^{\ell-1}$  is of size  $m_{e_{\ell-1}} \times m_{e_{\ell-1}}$  and  $M_e^{\ell,\ell-1}$  is of size  $m_{e_\ell} \times m_{e_{\ell-1}}$ .

2. Let  $L^\ell$  and  $L^{\ell-1}$  be the supermesh assembling matrices of the finite element spaces  $V^\ell$  and  $V^{\ell-1}$  respectively, mapping the local supermesh cell degrees of freedom to the global degrees of freedom of  $V^\ell$ . Assemble the local supermesh contributions together with

$$M = \left[ \begin{array}{c|c} (L^\ell)^T \text{diag}_e(M_e^\ell) L^\ell & (L^\ell)^T \text{diag}_e(M_e^{\ell,\ell-1}) L^{\ell-1} \\ \hline (L^{\ell-1})^T \text{diag}_e(M_e^{\ell,\ell-1})^T L^\ell & (L^{\ell-1})^T \text{diag}_e(M_e^{\ell-1}) L^{\ell-1} \end{array} \right]. \quad (4.44)$$

Observe that (4.44) and (4.12) agree since

$$\begin{aligned} M^s &= (L^s)^T \text{diag}_e(M_e^s) L^s, & \text{for } s \in \{\ell, \ell-1\}, \\ M^{\ell,\ell-1} &= (L^\ell)^T \text{diag}_e(M_e^{\ell,\ell-1}) L^{\ell-1}. \end{aligned} \quad (4.45)$$

Note that the above is again just the assembly of the contributions of each supermesh cell to the global mass matrices in matrix form. As we will see next, we actually do not need to assemble  $M$ , but only the local mass matrices  $M_e^\ell$  and  $M_e^{\ell,\ell-1}$  for each supermesh cell  $e$ .

### From global to local: the local coupling condition

We again use a divide-and-conquer strategy to split the global sampling problem into smaller local subproblems (cf. Remark 4.4). Suppose that we can sample a local Gaussian vector  $\mathbf{b}_e \sim \mathcal{N}(0, M_e)$  on each supermesh cell  $e$ . We can then separate  $\mathbf{b}_e$  into two Gaussian vectors  $\mathbf{b}_e^\ell$  and  $\mathbf{b}_e^{\ell-1}$  such that  $\mathbf{b}_e = [(\mathbf{b}_e^\ell)^T, (\mathbf{b}_e^{\ell-1})^T]^T$  and

$$\mathbf{b}_e^\ell \sim \mathcal{N}(0, M_e^\ell), \quad \mathbf{b}_e^{\ell-1} \sim \mathcal{N}(0, M_e^{\ell-1}), \quad \mathbb{E}[\mathbf{b}_e^\ell (\mathbf{b}_e^{\ell-1})^T] = M_e^{\ell,\ell-1}. \quad (4.46)$$

Since (4.46) is the local equivalent of (4.11), we refer to it as the *local coupling condition*. Finally, we can use the same approach as in (4.42) and assemble the coupled vectors  $\mathbf{b}^\ell$  and  $\mathbf{b}^{\ell-1}$  as

$$\mathbf{b}^\ell = \sum_{e=1}^n (L_e^\ell)^T \mathbf{b}_e^\ell, \quad \mathbf{b}^{\ell-1} = \sum_{e=1}^n (L_e^{\ell-1})^T \mathbf{b}_e^{\ell-1}, \quad (4.47)$$

where  $n$  is the number of supermesh cells. This enforces the correct distribution since sums of Gaussian random variables are Gaussian and the covariance structure is correct. In particular,

$$\begin{aligned} \mathbb{E}[\mathbf{b}^s (\mathbf{b}^s)^T] &= \sum_{i,j=1}^n (L_i^s)^T \mathbb{E}[\mathbf{b}_i^s (\mathbf{b}_j^s)^T] L_j^s = \sum_{i=1}^n (L_i^s)^T \mathbb{E}[\mathbf{b}_i^s (\mathbf{b}_i^s)^T] L_i^s \\ &= (L^s)^T \text{diag}_i(M_i^s) L^s = M^s, \quad \text{for } s \in \{\ell, \ell-1\}, \end{aligned} \quad (4.48)$$

and,

$$\begin{aligned}\mathbb{E}[\mathbf{b}^\ell(\mathbf{b}^{\ell-1})^T] &= \sum_{i,j=1}^n (L_i^\ell)^T \mathbb{E}[\mathbf{b}_i^\ell(\mathbf{b}_j^{\ell-1})^T] L_j^{\ell-1} = \sum_{i=1}^n (L_i^\ell)^T \mathbb{E}[\mathbf{b}_i^\ell(\mathbf{b}_i^{\ell-1})^T] L_i^{\ell-1} \\ &= (L^\ell)^T \text{diag}_i(M_i^{\ell,\ell-1}) L^{\ell-1} = M^{\ell,\ell-1},\end{aligned}\quad (4.49)$$

where we have used that  $\mathbf{b}_i^\ell$  and  $\mathbf{b}_j^\ell$  are independent for  $i \neq j$  for  $\ell \in \{\ell, \ell-1\}$  and that  $\mathbf{b}_i^\ell$  is independent from  $\mathbf{b}_j^{\ell-1}$  if  $i \neq j$ . Thus, again the global sampling problem can be recast as a series of much smaller, independent, local sampling problems.

Finally, it remains to devise a strategy for sampling realisations of the local vectors  $\mathbf{b}_e$  on a given supermesh cell  $e$ . The following result demonstrates that the covariance matrix of  $\mathbf{b}_e$  is singular and how such sampling can be simplified.

**Lemma 4.5.** *Let  $V^\ell$  and  $V^{\ell-1}$  be finite element spaces over two tessellations  $D_h^\ell$ ,  $D_h^{\ell-1}$  of the same domain. Let  $S_h$  be a supermesh of  $D_h^\ell$  and  $D_h^{\ell-1}$ . Let  $\phi_i^\ell$ ,  $\phi_j^{\ell-1}$  for  $i = 1, \dots, m_{e_\ell}$ ,  $j = 1, \dots, m_{e_{\ell-1}}$  be the basis functions of  $V^\ell$  and  $V^{\ell-1}$  respectively that have non-zero support over  $e$ . Let  $V^\ell|_e = \text{span}(\phi_1^\ell|_e, \dots, \phi_{m_{e_\ell}}^\ell|_e)$  and  $V^{\ell-1}|_e = \text{span}(\phi_1^{\ell-1}|_e, \dots, \phi_{m_{e_{\ell-1}}}^{\ell-1}|_e)$  be the restrictions of  $V^\ell$  and  $V^{\ell-1}$  to  $e$ . Assume that  $V^{\ell-1}|_e \subseteq V^\ell|_e$ , i.e. that the restrictions are nested, then*

$$\text{rank}(M_e) = \text{rank}(M_e^\ell) \quad \text{and} \quad M_e^{\ell-1} = (M_e^{\ell,\ell-1})^T (M_e^\ell)^{-1} M_e^{\ell,\ell-1}. \quad (4.50)$$

*Proof.* Since  $V^{\ell-1}|_e \subseteq V^\ell|_e$ , then we have that, for all  $j$ ,  $\phi_j^{\ell-1} \in V^\ell|_e$ , which in turn means that there exists a set of coefficients  $r_{ji} \in \mathbb{R}$  such that  $\phi_j^{\ell-1} = \sum_i r_{ji} \phi_i^\ell$ . Now, let  $R_e$  be a  $m_{e_\ell} \times m_{e_{\ell-1}}$  matrix such that  $(R_e)_{i,j} = r_{ji}$ , and define the vector functions

$$\boldsymbol{\phi}^{\ell-1} = \begin{bmatrix} \phi_1^{\ell-1} \\ \vdots \\ \phi_{m_{e_{\ell-1}}}^{\ell-1} \end{bmatrix}, \quad \boldsymbol{\phi}^\ell = \begin{bmatrix} \phi_1^\ell \\ \vdots \\ \phi_{m_{e_\ell}}^\ell \end{bmatrix}. \quad (4.51)$$

We then have that

$$\boldsymbol{\phi}^{\ell-1} = R_e^T \boldsymbol{\phi}^\ell. \quad (4.52)$$

This implies that we can now rewrite  $M_e^{\ell-1}$  as

$$M_e^{\ell-1} = \int_e \boldsymbol{\phi}^{\ell-1} (\boldsymbol{\phi}^{\ell-1})^T dx = R_e^T \int_e \boldsymbol{\phi}^\ell (\boldsymbol{\phi}^\ell)^T dx R_e = R_e^T M_e^\ell R_e, \quad (4.53)$$

since  $R_e$  is constant. Similarly, for  $M_e^{\ell,\ell-1}$  we have,

$$M_e^{\ell,\ell-1} = \int_e \boldsymbol{\phi}^\ell (\boldsymbol{\phi}^{\ell-1})^T dx = \int_e \boldsymbol{\phi}^\ell (\boldsymbol{\phi}^\ell)^T dx R_e = M_e^\ell R_e. \quad (4.54)$$

Combining (4.53) and (4.54) with the fact that  $M_e^\ell$  is invertible thus yields the second equation in (4.50), since

$$(M_e^{\ell,\ell-1})^T (M_e^\ell)^{-1} M_e^{\ell,\ell-1} = R_e^T M_e^\ell (M_e^\ell)^{-1} M_e^\ell R_e = R_e^T M_e^\ell R_e = M_e^{\ell-1}. \quad (4.55)$$

Pulling (4.43), (4.53) and (4.54) together we can now express  $M_e$  as,

$$M_e = \left[ \begin{array}{c|c} M_e^\ell & M_e^\ell R_e \\ \hline R_e^T M_e^\ell & R_e^T M_e^\ell R_e \end{array} \right] = \left[ \begin{array}{c|c} I & 0 \\ \hline R_e^T & I \end{array} \right] \left[ \begin{array}{c|c} M_e^\ell & 0 \\ \hline 0 & 0 \end{array} \right] \left[ \begin{array}{c|c} I & R_e \\ \hline 0 & I \end{array} \right], \quad (4.56)$$

where we have used the fact that  $M_e^\ell$  is symmetric. Since  $M_e$  is symmetric and the two block triangular matrices on the right-hand side of (4.56) are invertible, Sylvester's law of inertia [162] gives that

$$\text{rank}(M_e) = \text{rank} \left( \left[ \begin{array}{c|c} M_e^\ell & 0 \\ \hline 0 & 0 \end{array} \right] \right) = \text{rank}(M_e^\ell), \quad (4.57)$$

which concludes the proof.  $\square$

The assumptions of Lemma 4.5 are mild and are satisfied by most finite element families e.g. Lagrange elements.

Using Lemma 4.5 we can now sample  $\mathbf{b}_e = [(\mathbf{b}_e^\ell)^T, (\mathbf{b}_e^{\ell-1})^T]^T$  by enforcing the local coupling condition (4.46) as follows. For each supermesh cell  $e$ :

1. Compute  $M_e^{\ell,\ell-1}$  and the Cholesky factorisation  $M_e^\ell = H_e H_e^T$ .
2. Sample  $\mathbf{z}_e \sim \mathcal{N}(0, I)$  of length  $m_{e\ell}$  and set  $\mathbf{b}_e^\ell = H_e \mathbf{z}_e$ .
3. Compute  $\mathbf{b}_e^{\ell-1}$  as  $\mathbf{b}_e^{\ell-1} = (M^{\ell,\ell-1})^T H_e^{-T} \mathbf{z}_e$ .

Note that the  $\mathbf{b}_e^\ell$  and  $\mathbf{b}_e^{\ell-1}$  sampled this way satisfy the local coupling condition (4.46) since, by (4.50) and the fact that  $H_e$  is the Cholesky factor of  $M_e^\ell$ , we have that

$$\mathbb{E}[\mathbf{b}_e^\ell (\mathbf{b}_e^\ell)^T] = H_e \mathbb{E}[\mathbf{z}_e \mathbf{z}_e^T] H_e^T = M_e^\ell, \quad (4.58)$$

second,

$$\begin{aligned} \mathbb{E}[\mathbf{b}_e^{\ell-1} (\mathbf{b}_e^{\ell-1})^T] &= (M_e^{\ell,\ell-1})^T H_e^{-T} \mathbb{E}[\mathbf{z}_e \mathbf{z}_e^T] H_e^{-1} M_e^{\ell,\ell-1} \\ &= (M_e^{\ell,\ell-1})^T (H_e H_e^T)^{-1} M_e^{\ell,\ell-1} \\ &= (M_e^{\ell,\ell-1})^T (M_e^\ell)^{-1} M_e^{\ell,\ell-1} = M_e^{\ell-1}, \end{aligned} \quad (4.59)$$

and third,

$$\mathbb{E}[\mathbf{b}_e^\ell (\mathbf{b}_e^{\ell-1})^T] = H_e \mathbb{E}[\mathbf{z}_e \mathbf{z}_e^T] H_e^{-1} M_e^{\ell,\ell-1} = H_e H_e^{-1} M_e^{\ell,\ell-1} = M_e^{\ell,\ell-1}. \quad (4.60)$$

In the case in which the transformation to the reference cell is affine (such as with Lagrange elements on simplices) the sampling can be made more efficient by sampling white noise directly on the supermesh and then interpolating it onto the parent meshes. This strategy exploits the following result.

**Corollary 4.6** (of Lemma 4.5). *Let  $V^\ell$  and  $V^{\ell-1}$  be FEM approximation subspaces over two triangulations  $D_h^\ell$ ,  $D_h^{\ell-1}$  of the same domain. Let  $S_h$  be a supermesh of  $D_h^\ell$  and  $D_h^{\ell-1}$  and let  $V^S$  be a FEM approximation subspace over  $S_h$ . With the same notation as in Lemma 4.5, for each supermesh cell  $e$  let  $V^S|_e$ ,  $V^\ell|_e$  and  $V_e^{\ell-1}$  be the restrictions of  $V^S$ ,  $V^\ell$  and  $V^{\ell-1}$  to  $e$ . Let  $M_e^S$  be the local mass matrix over  $V^S|_e$ . Assume that  $V^\ell|_e, V^{\ell-1}|_e \subseteq V^S|_e$ , i.e. that the parent mesh restrictions are nested within the supermesh restriction. Then there exist local interpolation matrices  $(R_e^\ell)^T$  and  $(R_e^{\ell-1})^T$  such that*

$$M_e^\ell = (R_e^\ell)^T M_e^S R_e^\ell, \quad M_e^{\ell-1} = (R_e^{\ell-1})^T M_e^S R_e^{\ell-1}, \quad M_e^{\ell,\ell-1} = (R_e^\ell)^T M_e^S R_e^{\ell-1}. \quad (4.61)$$

*Proof.* Let  $l \in \{\ell, \ell - 1\}$ . The proof for the first two equations in (4.61) follows from the first part of the proof of Lemma 4.5 by replacing  $\ell - 1$  with  $l$  and  $\ell$  with  $S$ . This argument gives us that

$$\boldsymbol{\phi}^l = (R_e^l)^T \boldsymbol{\phi}^S, \quad (4.62)$$

from which we also obtain the last relation in (4.61) since

$$M_e^{\ell,\ell-1} = \int_e \boldsymbol{\phi}^\ell (\boldsymbol{\phi}^{\ell-1})^T dx = (R_e^\ell)^T \int_e \boldsymbol{\phi}^S (\boldsymbol{\phi}^S)^T dx R_e^{\ell-1} = (R_e^\ell)^T M_e^S R_e^{\ell-1}. \quad (4.63)$$

□

By using this result and the strategy highlighted in Remark 4.5, we can sample  $\mathbf{b}_e^l$  for  $l \in \{\ell, \ell - 1\}$  by computing

$$\mathbf{b}_e^l = (R_e^l)^T H_r |e_r|^{-1/2} |e|^{1/2} \mathbf{z}_e, \quad \text{with } \mathbf{z}_e \sim \mathcal{N}(0, I), \quad (4.64)$$

since Remark 4.5 yields the relation  $M_e^S / |e| = H_r H_r^T / |e_r| = \text{const}$ , where  $H_r$  is the Cholesky factor of the local mass matrix over the reference cell  $e_r$ . Note that  $H_r$  has to be computed only once. The advantage of performing this operation is that it avoids the assembly and factorisation of each supermesh cell local mass matrix.

*Remark 4.6* (Simpler cases: nested meshes and  $p$ -refinement). In the case in which the meshes of the MLMC hierarchy are nested, everything discussed is still valid by taking the supermesh to be the finer of the two meshes that define the MLMC level. In the case in which the MLMC hierarchy is constructed by using  $p$ -refinement there is only one mesh in the hierarchy and everything still applies by taking this mesh to be the ‘supermesh’. In both cases a supermesh construction is not required in practice.

*Remark 4.7.* The coupling approach presented can also be used to couple the same white noise sample over the whole hierarchy of meshes. This enables the use of geometric full-multigrid [169] to solve the problem given by (4.7)–(4.8) on the finer grid with optimal multigrid complexity.

*Remark 4.8.* The sampling approaches just presented have linear cost complexity in the number of mesh cells (in the independent case) and of supermesh cells (in the coupled realisation case). In turn, thanks to Corollary 4.4, the sampling strategy in the coupled case is linear in the number of cells of the finer of the parent meshes.

## 4.5 Numerical results

In this section we investigate the performance of the techniques presented. We consider the following PDE:

$$\begin{aligned} -\nabla \cdot (e^{u(\mathbf{x}, \omega)} \nabla q(\mathbf{x}, \omega)) &= 1, & \mathbf{x} \in G = (-0.5, 0.5)^d, & \omega \in \Omega, \\ q(\mathbf{x}, \omega) &= 0, & \mathbf{x} \in \partial G, & \omega \in \Omega, \end{aligned} \quad (4.65)$$

where  $u$  is a Matérn field as given by (2.27) with mean and variance chosen so that  $e^{u(\mathbf{x}, \omega)}$  has mean 1 and standard deviation 0.2. We choose  $D = (-1, 1)^d$  as the outer computational domain on which to solve (4.2). The output functional of interest we consider here is the  $L^2(G)$  norm of  $q$  squared, namely  $P(\omega) = \|q\|_{L^2(G)}^2(\omega)$ .

We approximate the coefficient  $u$  in (4.65) by solving (4.2) using the FEM. We solve (4.2) and (4.65) with the FEniCS software package [120] and we discretise the two problems by using continuous Lagrange finite elements of the same degree. For the linear solver, we use the BoomerAMG algebraic multigrid algorithm from Hypre [57] as a preconditioner and the conjugate gradient routine of PETSc [15] for all equations. As convergence criterion for the solver we require that the absolute size of the preconditioned residual norm is below a tolerance of  $10^{-10}$ . We use the libsupermesh software package [121] for the supermesh constructions.

*Remark 4.9.* In our work we did not investigate the effect of the quadrature error arising from the FEM linear system assembly. However, the FEniCS software automatically chooses enough quadrature nodes so that all the integrals are computed exactly given the polynomial degree of the FEM approximation subspaces involved. This is of course more costly than it should be, but makes all of the numerical results in this chapter free from quadrature error. We leave the analysis of quadrature error for future research.

When using  $h$ -refinement, we construct the MLMC mesh hierarchies  $\{D_h^\ell\}_{\ell=1}^L$  and  $\{G_h^\ell\}_{\ell=1}^L$  in such a way that  $G_h^\ell$  is embedded within  $D_h^\ell$  for all  $\ell$ , but  $D_h^{\ell-1}$  and  $G_h^{\ell-1}$  are not nested respectively within  $D_h^\ell$  and  $G_h^\ell$  for all  $\ell > 1$ . As the meshes are non-nested, a supermesh construction is required to couple each MLMC level. The mesh hierarchies we use are composed of  $L = 9$  meshes in 2D and  $L = 5$  meshes in 3D. The coarsest mesh in each hierarchy is uniform, while the other meshes are non-uniform and unstructured. We take our sampling domain to be  $(-1, 1)^d$ . This, and the values of the Matérn parameters chosen in this section ensures that the error in the covariance due to domain truncation is below machine precision in 2D and around  $10^{-15}$  in 3D, cf. [103]. Since the convergence behaviour of the FEM is dependent on the quality of the mesh used, we try to sanitise our numerical results from this effect by choosing meshes whose quality indicators do not vary excessively throughout the hierarchies. Basic properties of the different meshes and number of cells of the constructed supermeshes are summarised in Tables 4.1 and 4.2. Note that, as stated by Corollary 4.4, the number of cells in the supermesh is in practice always bounded by a constant times the number of cells of the finer parent mesh. This constant is dimension-dependent, and larger in 3D than 2D (cf. table 4.2). We do not compare

$\ell$ (2D)	$h_\ell$	$n_\ell$	$(RR_{\min}, RR_{\max})$	$n_{S_\ell}/n_\ell$
1	0.707	32	(0.83, 0.83)	n/a
2	0.416	120	(0.61, 1)	2.03
3	0.194	500	(0.61, 1)	2.32
4	0.098	2106	(0.55, 1)	2.45
5	0.049	8468	(0.45, 1)	2.44
6	0.024	33686	(0.46, 1)	2.46
7	0.012	134170	(0.41, 1)	2.46
8	0.006	535350	(0.42, 1)	2.46
9	0.003	2143162	(0.42, 1)	2.47

Table 4.1: Properties of the 2D mesh hierarchy: mesh level  $\ell$ , maximal cell size  $h_\ell$ , number of cells  $n_\ell$ , minimal and maximal cell radius ratios  $RR_{\min}$  and  $RR_{\max}$ , and the number of cells of the supermesh constructed using the meshes on levels  $\ell$  and  $\ell - 1$  as parent meshes  $n_{S_\ell}$ .  $RR$  is computed as  $d \times r_{in}^e / r_{circ}^e$ , where  $r_{in}^e$  and  $r_{circ}^e$  are the in-radius and the circumradius of cell  $e$  respectively. Note that the cell size roughly decreases proportional to  $2^{-\ell}$ .

$\ell$ (3D)	$h_\ell$	$n_\ell$	$(RR_{\min}, RR_{\max})$	$(DA_{\min}, DA_{\max})$	$n_{S_\ell}/n_\ell$
1	0.866	384	(0.72, 0.72)	(0.79, 1.57)	n/a
2	0.437	7141	(0.22, 1)	(0.21, 2.82)	17
3	0.280	22616	(0.18, 1)	(0.21, 2.83)	66
4	0.138	190081	(0.13, 1)	(0.21, 2.85)	42
5	0.070	1519884	(0.12, 1)	(0.21, 2.85)	46

Table 4.2: Properties of the 3D mesh hierarchy: mesh level  $\ell$ , maximal cell size  $h_\ell$ , number of cells  $n_\ell$ , minimal and maximal cell radius ratios  $RR_{\min}$  and  $RR_{\max}$ , the minimum and maximum cell dihedral angles  $DA_{\min}$  and  $DA_{\max}$  respectively, and the number of cells of the supermesh constructed using the meshes on levels  $\ell$  and  $\ell - 1$  as parent meshes. Note that the cell size of the last three levels roughly decreases proportional to  $2^{-\ell}$ .

this constant with the quasi-uniformity constants here as the constant in the bound given by Theorem 4.3 is not sharp.

In Tables 4.3 and 4.4 we present some representative average CPU timings over  $N = 1000$  realisations of coupled white noise and Matérn field samples in 2D and 3D over  $(-1, 1)^d$  and with  $\nu = 2 - d/2$ . These timings have been obtained in serial by using a single thread on a computing node with an Intel® Xeon® Gold 6140 CPU (2.30GHz). We consider both uniform structured, and non-nested unstructured hierarchies, but we do not exploit the structuredness of the first hierarchy in the implementation. These timings do not include offline one-off costs such as the assembly of the coupled linear systems deriving from the discretisation of Equations (4.7) and (4.8). We note that in 2D and in the nested case in 3D the cost of sampling

white noise realisations is always lower than the cost of solving the coupled linear system. However, in the non-nested case in 3D, the sampling of white noise becomes more expensive (by a factor of 20 approximately), mainly due to the larger number of supermesh cells, cf. Table 4.2. However, there is most likely room for improvement since our implementation of the white noise sampling routine is currently Python-based (while the solver uses PETSc routines which are heavily optimised). We leave a more efficient implementation to future work.

*Remark 4.10.* The MLMC method presented is designed to work with non-nested unstructured grid hierarchies. In this setting, a supermesh construction might be unavoidable even in the case in which the field is sampled via an alternative method such as e.g. circulant embedding [49], since this is the only way in which the sampled field can be integrated exactly on the target non-nested unstructured grid. For low smoothness fields ( $\nu \leq 1$ ), Graham et al. have proved that there is no loss in the convergence rate due to non-nested interpolation [74]. However, this operation, albeit faster, still introduces extra bias, and it is likely to still harm convergence when smoother fields are used. All in all, we remark that the high white noise sampling costs observed in the non-nested 3D case stem directly from the supermesh construction, and that similar timings would also be observed if a supermesh construction were used to integrate a field sampled on a structured grid over a target non-nested mesh.

$\ell$	2	3	4	5	6	7	8	9
$\dot{W}$ sample (NU)	$2 \times 10^{-4}$	$3 \times 10^{-4}$	$4 \times 10^{-4}$	$8 \times 10^{-4}$	$2 \times 10^{-3}$	$8 \times 10^{-3}$	0.004	0.25
$\dot{W}$ sample (NN)	$5 \times 10^{-4}$	$6 \times 10^{-4}$	$2 \times 10^{-3}$	$8 \times 10^{-3}$	0.031	0.130	0.508	2.09
solver (NU)	$5 \times 10^{-4}$	$6 \times 10^{-4}$	$1 \times 10^{-3}$	$5 \times 10^{-3}$	0.017	0.087	0.421	1.86
solver (NN)	$6 \times 10^{-4}$	$9 \times 10^{-4}$	$3 \times 10^{-3}$	0.010	0.046	0.212	0.947	3.98

Table 4.3: Average timings over 1000 realisations for the coupled sampling of coupled white noise and Matérn fields ( $\nu = 1$ ) for MLMC in the nested uniform (NU) and non-nested (NN) case in 2D. The timings do not include offline one-off costs, such as the set-up of the linear solver. The meshes used for the non-nested case are the same as in Table 4.1, the meshes used for the nested case are structured uniform grids of the domain  $(-1, 1)^2$  with  $2^{2\ell+5}$  triangles per grid. We note that the cost of sampling white noise is always lower than the cost of solving (4.7) and (4.8).

When using  $p$ -refinement, we define the MLMC levels by taking the coarsest mesh in the 2D hierarchy and by increasing the polynomial degree of the FEM subspaces linearly so that  $p_\ell = \ell$  for  $\ell = 1, \dots, L$ , with  $L = 9$ . We do not consider  $p$ -refinement in the 3D case as it does not offer any additional complications other than an increased computational cost.

#### 4.5.1 Matérn field convergence

We first address the convergence of the solution of (4.2) to the Matérn field of interest. For this purpose, we fix  $E[u] = 0$  and  $\sigma^2 = 1$ . In practice, the exact solution

$\ell$	1	2	3	4	5
W sample (NU)	$3 \times 10^{-4}$	$5 \times 10^{-4}$	$2 \times 10^{-3}$	0.013	0.12
W sample (NN)	$4 \times 10^{-4}$	0.06	0.56	4.3	34.1
solver (NU)	$4 \times 10^{-4}$	$1 \times 10^{-3}$	$9 \times 10^{-3}$	0.078	0.63
solver (NN)	$4 \times 10^{-4}$	$4 \times 10^{-3}$	0.027	0.23	1.85

Table 4.4: Average timings over 1000 realisations for the coupled sampling of coupled white noise and Matérn fields ( $\nu = 1/2$ ) for MLMC in the nested uniform (NU) and non-nested (NN) case in 3D. The timings do not include offline one-off costs, such as the set-up of the linear solver. The meshes used for the non-nested case are the same as in Table 4.2, the meshes used for the nested case are structured uniform grids of the domain  $(-1, 1)^3$  with  $5\sqrt{3} \times 2^{3\ell}$  triangles per grid.

of (4.2) is not known, so we consider the coupled equations (4.7) and (4.8) instead. We monitor the quantities

$$\left| \mathbb{E} \left[ \|u\|_{L^2(G)}^2 - \|u_{\ell-1}\|_{L^2(G)}^2 \right] \right|, \quad \mathbb{V} \left[ \|u_\ell\|_{L^2(G)}^2 - \|u_{\ell-1}\|_{L^2(G)}^2 \right]. \quad (4.66)$$

Note that the value of  $\mathbb{E}[\|u\|_{L^2(G)}^2]$  is known up to the error introduced by truncating  $\mathbb{R}^d$  to  $D$  since we can exchange the order of expectation and integration:

$$\mathbb{E} \left[ \|u\|_{L^2(G)}^2 \right] = \mathbb{E} \left[ \int_G u^2 \, d\mathbf{x} \right] = \int_G \mathbb{E}[u^2] \, d\mathbf{x} \approx \sigma^2 |G|, \quad (4.67)$$

where we have used the fact that  $\mathbb{E}[u^2] \approx \sigma^2$  for all  $\mathbf{x} \in G$  (the relation only holds approximately due to domain truncation error).

We derived theoretical estimates for the expected strong error convergence rates in Chapter 3. The quantity on the left in (4.66) is called the weak error and its convergence rate generally twice that of the strong error (see [28] for the case  $k \leq 1$ ), provided that the polynomial degree of the FEM basis is large enough. We are not aware of any error estimates in the literature for the variance in (4.66), but the convergence order observed in practice is usually twice that of the expectation (see for Example [41]), provided that the polynomial degree of the FEM basis is sufficiently high.

We consider the convergence behaviour of the FEM approximation of the solution of (4.1) in the  $h$ -refinement case with the sampling strategy described in Section 4.4. We fix  $\lambda = 0.2$  and we consider Matérn fields of smoothness  $\nu = 1$ ,  $\nu = 3$  ( $k = 1$  and  $k = 2$  respectively in 2D) and  $\nu = 1/2$  ( $k = 1$  in 3D). For the  $\nu = 1$  and  $\nu = 1/2$  cases we use continuous piecewise linear (P1) elements, while for the  $\nu = 3$  case we use continuous piecewise quadratic (P2) elements.

Since each sample drawn by solving (4.1) is computationally expensive we are unable to take large numbers of samples as is generally done for 1D stochastic differential equations [68]. We therefore take  $N_\ell = 5000$  Monte Carlo samples on all levels in 2D and  $N_\ell = 1000$  samples in 3D. To verify that these numbers of samples are sufficient for accurate representation, we compute approximate 99.73% confidence

intervals (CIs) for all the quantities of interest as  $3\bar{\sigma}_\ell/\sqrt{N_\ell}$ , where  $\bar{\sigma}_\ell$  is the sample standard deviation of the output functional of interest on level  $\ell$ . In all cases considered here but one, the FEM error dominates and the confidence intervals are negligibly small (so small that they would not be visible on the convergence plots). The relatively small number of samples only becomes a problem in the  $\nu = 3$  case where the FEM convergence is much faster and the Monte Carlo error dominates. In this case we replace the  $\|u\|_{L^2(G)}$  term in the expectation in (4.66) with  $\|u_\ell\|_{L^2(G)}$ , and we instead monitor the convergence of the following quantity,

$$\left| \mathbb{E} \left[ \|u_\ell\|_{L^2(G)}^2 - \|u_{\ell-1}\|_{L^2(G)}^2 \right] \right|. \quad (4.68)$$

The advantage of doing this is that the variance of this error measure decreases with the level (see Figure 4.2) and 5000 samples are enough to obtain good accuracy.

Results are shown in Figures 4.2 (2D) and 4.3 (3D). For both the 2D and 3D experiments, we observe the theoretically predicted convergence rates in terms of the mesh size (after a pre-asymptotic regime). However, we note how convergence is less regular than expected (especially in the 3D case) because of the unstructured meshes employed. This behaviour does not appear when uniform meshes are used (not shown). Apart from the  $\nu = 3$  case, the convergence order of the variance seems to be twice the convergence order of the expectation. In the  $\nu = 3$  case we observe order 6 for the variance with P2 elements (Figure 4.2) and order 8 with P3 elements (not shown). We conjecture that the variance convergence order is bounded by  $2(p+1)$ , where  $p$  is the polynomial degree of the FEM basis functions.

In Figure 4.4, we compare the covariances of the coupled Matérn fields obtained by solving (4.7) and (4.8) on the finest level of the MLMC hierarchy with the exact Matérn covariance given by (2.27). The estimated covariances match each other and the exact expression closely, demonstrating that our coupling technique is accurate also in practice.

As a final verification step, we check that the coupled fields are consistent with the telescoping sum in (2.15), i.e. if we let  $a$ ,  $b$ ,  $c$  be the MC approximations of  $\mathbb{E}[\|u_\ell\|_{L^2(G)}^2 - \|u_{\ell-1}\|_{L^2(G)}^2]$ ,  $\mathbb{E}[\|u_\ell\|_{L^2(G)}^2]$  and  $\mathbb{E}[\|u_{\ell-1}\|_{L^2(G)}^2]$  respectively, we aim to verify that

$$a - b + c \approx 0, \quad (4.69)$$

at least to within the Monte Carlo accuracy. In Figure 4.5, we plot the quantity

$$T(a, b, c) \equiv \frac{|a - b + c|}{3(\sqrt{\mathbb{V}_a} + \sqrt{\mathbb{V}_b} + \sqrt{\mathbb{V}_c})}, \quad (4.70)$$

for different levels and Matérn smoothness parameters  $\nu$ , where  $\mathbb{V}_a$ ,  $\mathbb{V}_b$  and  $\mathbb{V}_c$  are the Monte Carlo approximations of the variances of  $a$ ,  $b$  and  $c$ . The probability of this ratio  $T$  being greater than 1 is less than 0.3% (for further details, see [68]). We observe that  $T$  ranges between 0 and 0.4 for the levels and smoothness parameters tested (Figure 4.5), and in particular is well below 1. This indicates that our implementation of the MLMC algorithm correctly satisfies the telescoping summation formulation.

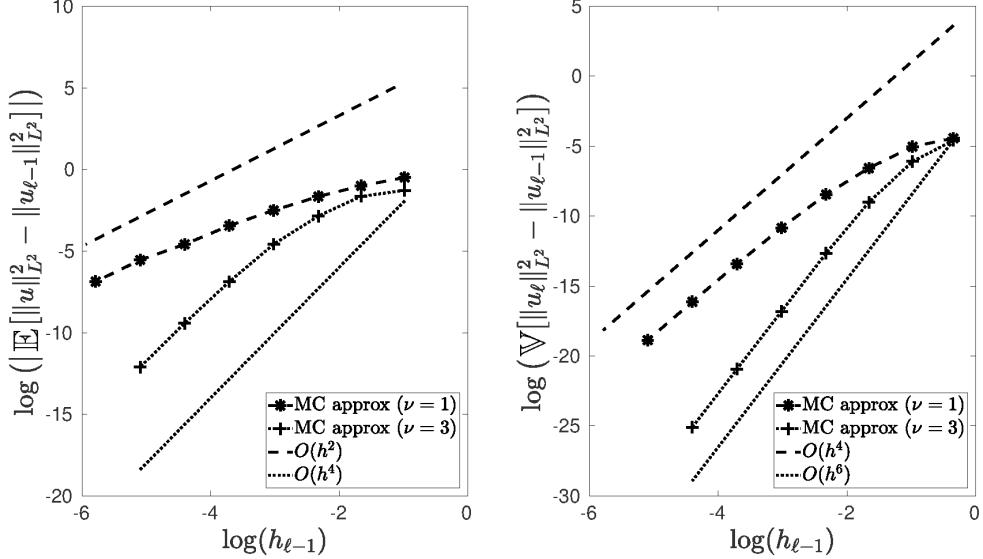


Figure 4.2: Convergence behaviour of the FEM approximation to (4.1) with  $h$ -refinement in 2D. Plots show (the natural logarithm of) the expected value  $\mathbb{E}$  (left) and variance  $\mathbb{V}$  (right) versus maximal mesh size  $h_\ell$ . For each level  $\ell$ , the fields  $u_\ell$  and  $u_{\ell-1}$  have been sampled by coupling white noise realisations as described in Section 4.4. As mentioned in the text, to compute the expected value in the  $\nu = 3$  case we have replaced  $\|u\|_{L^2(G)}$  with  $\|u_\ell\|_{L^2(G)}$ .

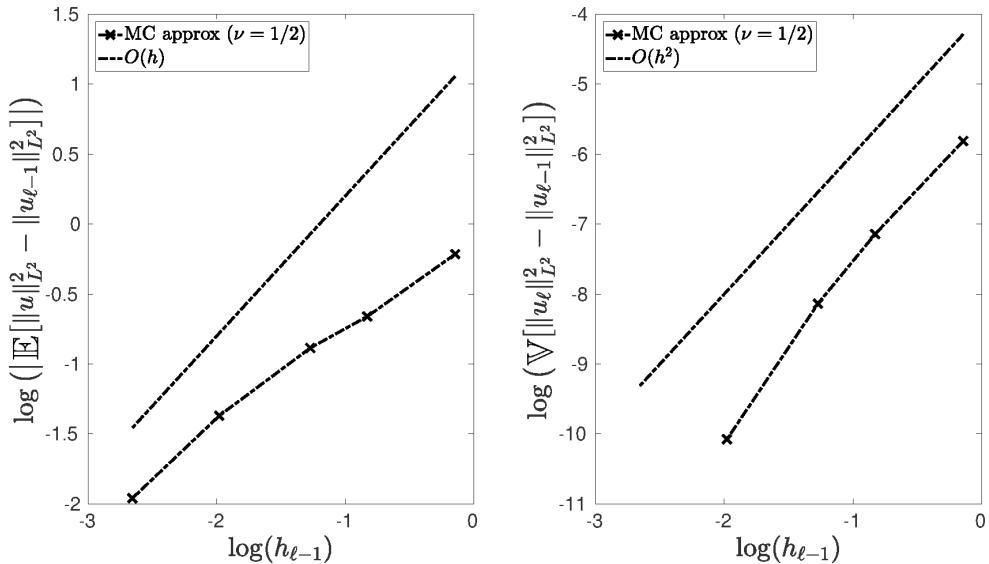


Figure 4.3: Convergence behaviour of the FEM approximation to (4.1) with  $h$ -refinement in 3D. Plots show (the natural logarithm of) the expected value  $\mathbb{E}$  (left) and variance  $\mathbb{V}$  (right) versus maximal mesh size  $h_l$ . The fields  $u_\ell$  and  $u_{\ell-1}$  have been sampled by coupling white noise realisations as described in Section 4.4.

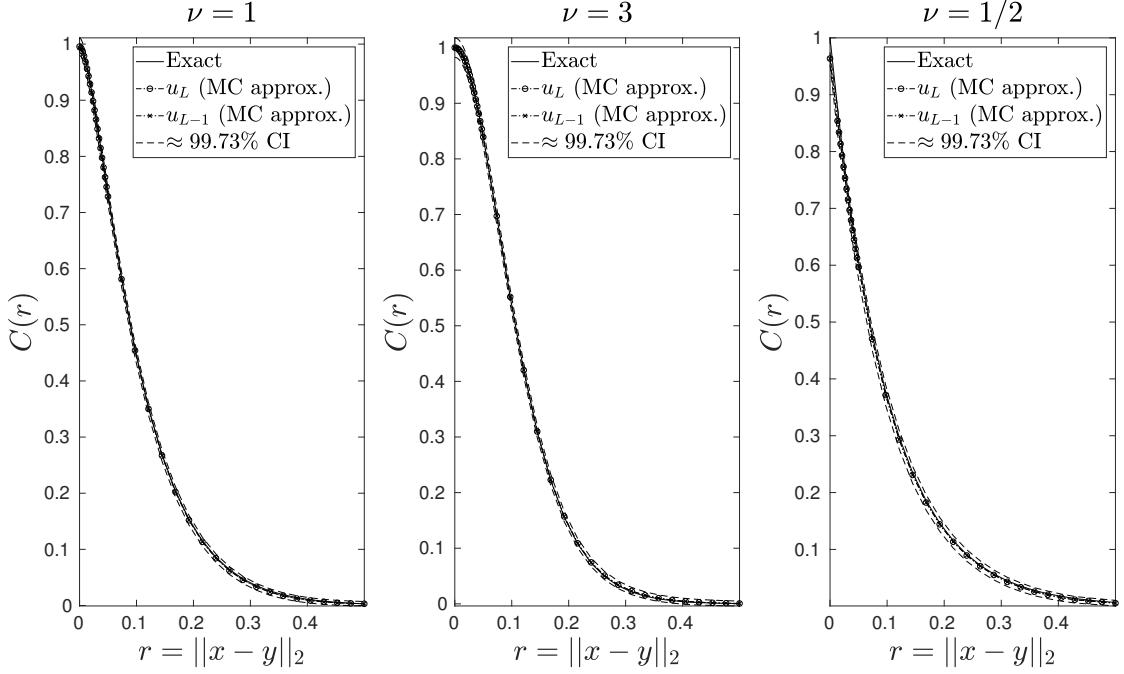


Figure 4.4: Plot of exact covariances and sample covariances vs distance  $r$  of the FEM solutions of (4.7) and (4.8) for three different values of  $\nu$  in the  $h$ -refinement case. The exact covariance  $C(r)$  is given by (2.27). For the  $\nu = 3$  case, an extra elliptic PDE solve is needed, see (2.43).

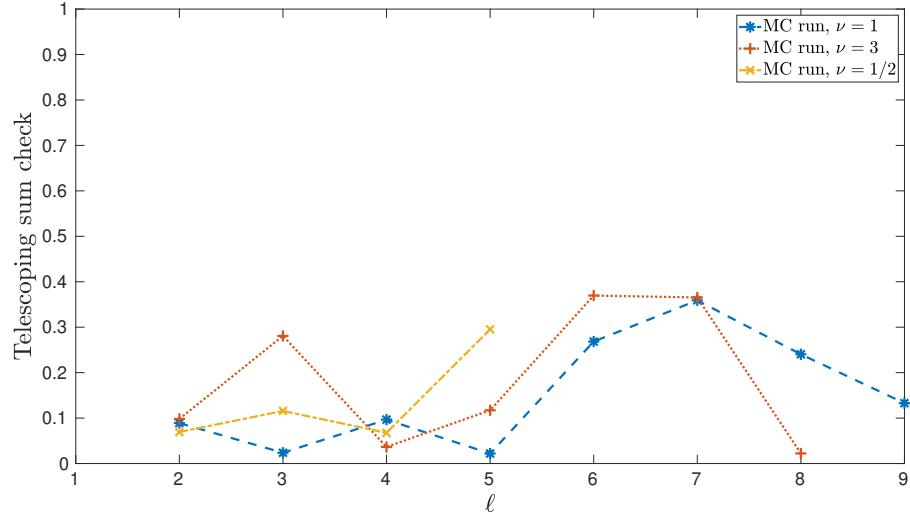


Figure 4.5: Telescoping sum consistency check. Plot of  $T(a, b, c)$  as defined by (4.70) versus level  $l$  for  $a = \mathbb{E}[\|u_\ell\|_{L^2(G)}^2 - \|u_{\ell-1}\|_{L^2(G)}^2]$ ,  $b = \mathbb{E}[\|u_\ell\|_{L^2(G)}^2]$  and  $c = \mathbb{E}[\|u_{\ell-1}\|_{L^2(G)}^2]$  for different smoothness parameters  $\nu$ .

### 4.5.2 MLMC convergence

We now consider the convergence of the multilevel Monte Carlo method applied to (4.65). In the case where  $u$  is sampled exactly, the assumptions of the MLMC convergence theorem (Theorem 2.1) hold for the  $h$ -refinement case with constants  $\alpha = 2$  and  $\beta = 4$  [39]. Furthermore, since we use multigrid to solve (4.65) and (4.2) we have  $\gamma = d$ . In the case where  $\nu > 1$ , the Matérn field smoothness increases [3] and we expect higher convergence rates for the solution of (4.65). For integer  $\nu$  and exact sampling of  $u$ , if the domain  $G$  is of class  $C^{\nu+1}$ , then the MLMC parameter values are given by  $\alpha = \min(\nu + 1, p + 1)$  and  $\beta = 2\alpha$ , where  $p$  is the polynomial degree of the Lagrange elements used [115].

In our case,  $u$  is approximated with the FEM and this could affect convergence. To verify that this is not what happens in practice, we first solve (4.1) with FEM for the same parameter values as in SubSection 4.5.1, namely  $\lambda = 0.2$ ,  $\nu = 1$  and  $\nu = 3$  ( $k = 1$  and  $k = 2$  respectively in 2D) and  $\nu = 1/2$  ( $k = 1$  in 3D) using P1 elements for  $\nu = 1/2$  and  $\nu = 1$  and P2 elements for  $\nu = 3$ . We then use the approximated Matérn fields computed this way as coefficients in (4.65), which we solve again using the same choice of finite elements.

Results are shown in Figures 4.6 and 4.7. We observe that the convergence is unaffected by the approximation of the Matérn fields and that the estimated convergence orders agree with the theory [39] apart from some discrepancies in the 3D case. This irregular behaviour is probably due to the non-uniformity of the hierarchy (as we see from Table 4.2, the quality of the 3D meshes decreases with the level). This issue does not arise if the same numerical experiment is performed using a uniform hierarchy instead (hierarchy mesh sizes given by  $h_\ell = 1.732 \times 2^{-\ell}$ , see Figure 4.8).

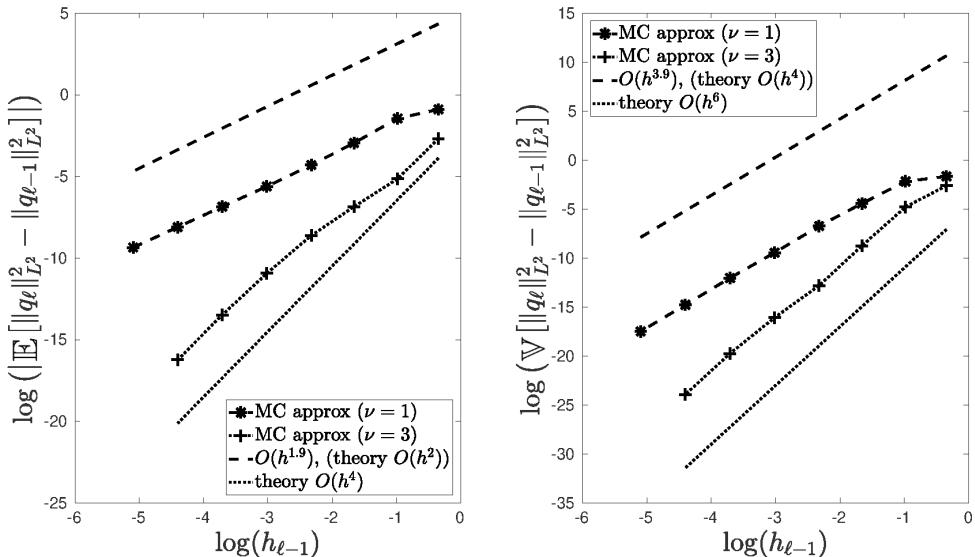


Figure 4.6: Convergence behaviour of the FEM approximation to the solution of (4.65) with  $h$ -refinement in 2D. The estimated convergence orders agree with the theory [39, 115].

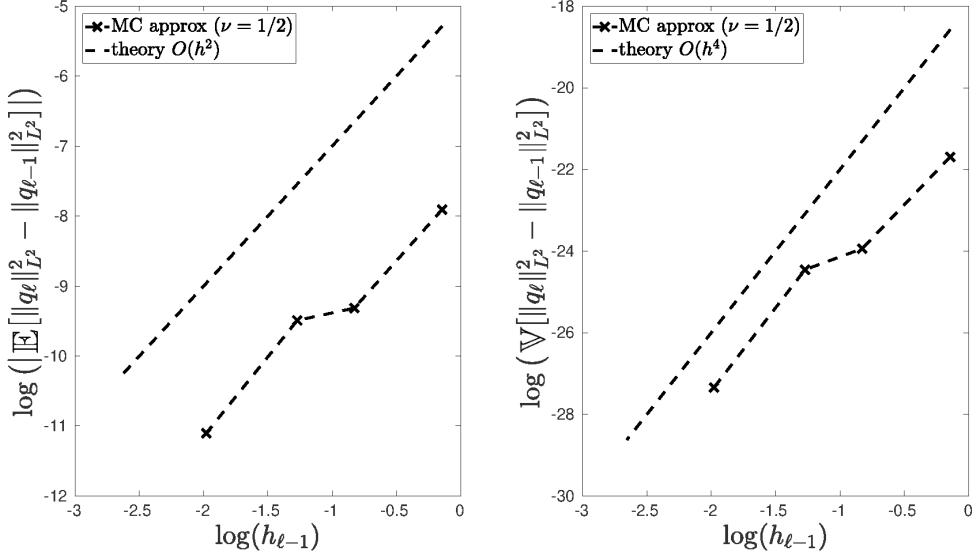


Figure 4.7: Convergence behaviour of the FEM approximation to the solution of (4.65) with  $h$ -refinement in 3D.

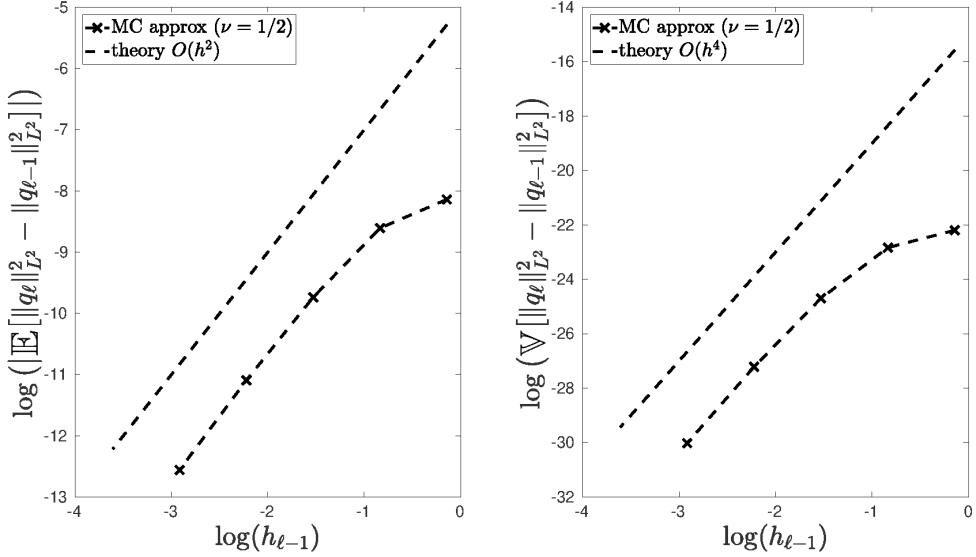


Figure 4.8: Convergence behaviour of the FEM approximation to the solution of (4.65) with  $h$ -refinement in 3D using a hierarchy of uniform meshes. The estimated convergence orders agree with the theory [39]. The mesh sizes are given by  $h_\ell = 1.732 \times 2^{-\ell}$ .

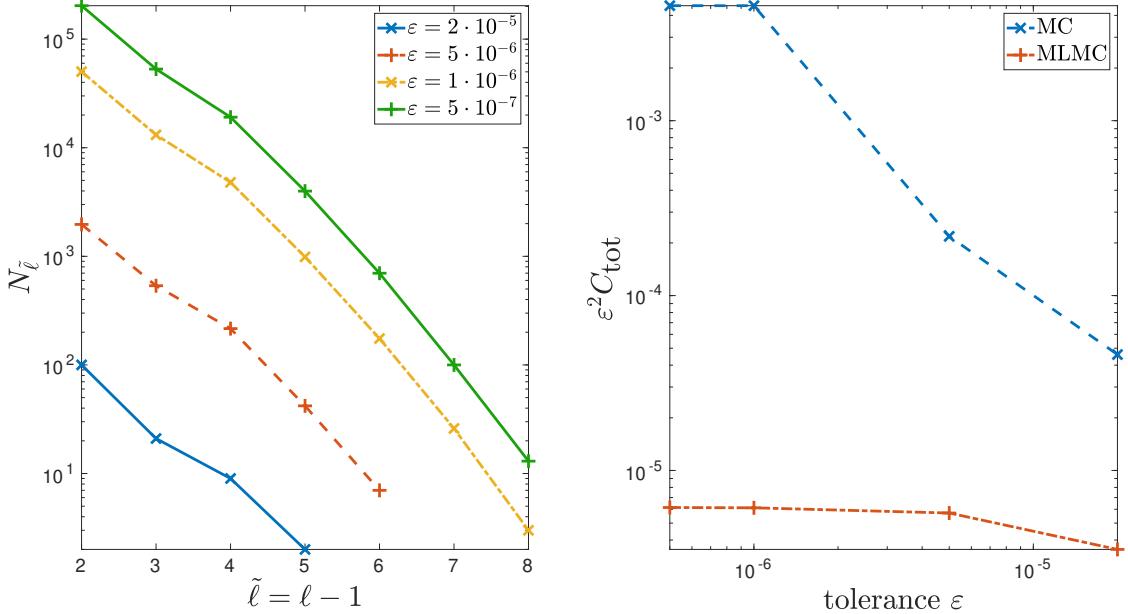


Figure 4.9: MLMC convergence for the solution of (4.65). In the plot on the left we show how the MLMC algorithm automatically selects the optimal number of samples  $N_{\tilde{\ell}}$  on each level to achieve a given tolerance  $\varepsilon$ . Note that the MLMC routine uses the second mesh in the hierarchy described in Table 4.1 to define the first level  $\tilde{\ell}$ . The first mesh in Table 4.1 is dropped since it is too coarse and it would not bring any significant advantage to the performance of MLMC [68]. In the plot on the right we compare the efficiency of MLMC with standard MC for different tolerances. The savings of MLMC with respect to standard Monte Carlo are considerable.

We now investigate how MLMC performs in practice. We use standard Monte Carlo and MLMC to estimate  $\mathbb{E}[\|q\|_{L^2(G)}]$  at the same accuracy for  $\nu = 1$  (2D), P1 elements and for different mean square error tolerances  $\varepsilon$ . Again, the coefficient  $u$  of (4.65) is also approximated with the FEM. We keep track of the total computational cost  $C_{\text{tot}}$  and, in the MLMC case, of the number of samples  $N_{\ell}$  taken on each level.

Results are shown in Figure 4.9. We observe that the number of levels used increases as the tolerance  $\varepsilon$  decreases (Figure 4.9, left). This behaviour reflects the targeted weak error accuracy [68]: the number of samples is chosen by the MLMC algorithm so as to optimise the total computational effort [69] and it decreases with the level, with many samples on the coarse levels and only a few on the fine levels. As  $\beta > \gamma$  (cf. Theorem 2.1), we expect the total cost of the MLMC algorithm  $C_{\text{tot}}$  to be proportional to  $\varepsilon^{-2}$ . Figure 4.9 (right) shows  $\varepsilon^2 C_{\text{tot}}$  versus  $\varepsilon$ , where  $C_{\text{tot}}$  is computed using pseudo-costs as  $\sum_{\ell=1}^L 2^{\gamma\ell} N_{\ell}$ . Indeed, we observe a near constant  $\varepsilon^2 C_{\text{tot}}$  for the MLMC algorithm across multiple choices of  $\varepsilon$ . Figure 4.9 also compares the MLMC cost with the cost of obtaining an estimate of the same accuracy with standard Monte Carlo. We observe that the MLMC algorithm offers significant computational savings compared to standard Monte Carlo, with an improvement in the total cost  $C_{\text{tot}}$  of up to 3 orders of magnitude (Figure 4.9, right).

Finally, we consider the convergence of MLMC with  $p$ -refinement. We follow the

same procedure as in the  $h$ -refinement case and solve (4.65) after approximating the coefficient  $u$  by solving (4.1) with FEM. This time, however, we fix the mesh to be the coarsest mesh in the 2D hierarchy (cf. Table 4.1) and we consider a hierarchy of continuous piecewise polynomial elements of increasing polynomial degree  $p = 1, \dots, 8$ . We investigate the convergence behaviour for different values of  $\nu$ , namely  $\nu \in \{1, 7, 31\}$  (corresponding to  $k \in \{1, 3, 15\}$ ).

We observe in Figure 4.10 that convergence is geometric (the error decreases exponentially as the polynomial degree  $p$  grows). The solution of (4.65) is actually almost surely not analytic and we would therefore expect algebraic convergence (i.e. the error decreases polynomially as  $p$  grows), cf. Chapter 3 and [12]. We hypothesise that this better-than-expected convergence is in fact pre-asymptotic behaviour and that the geometric convergence will eventually plateau and switch to a slower algebraic rate that depends on the smoothness of  $u$  (the larger  $\nu$ , the faster the convergence) [80, 152]. However, apart from the  $\nu = 1$  case for which the convergence plot begins to tail off, this is not observed for the polynomial degrees considered. We note that the larger the smoothness parameter  $\nu$  is, the faster the expected value converges. The variance convergence order, on the other hand, seems to be unaffected by the value of  $\nu$ .

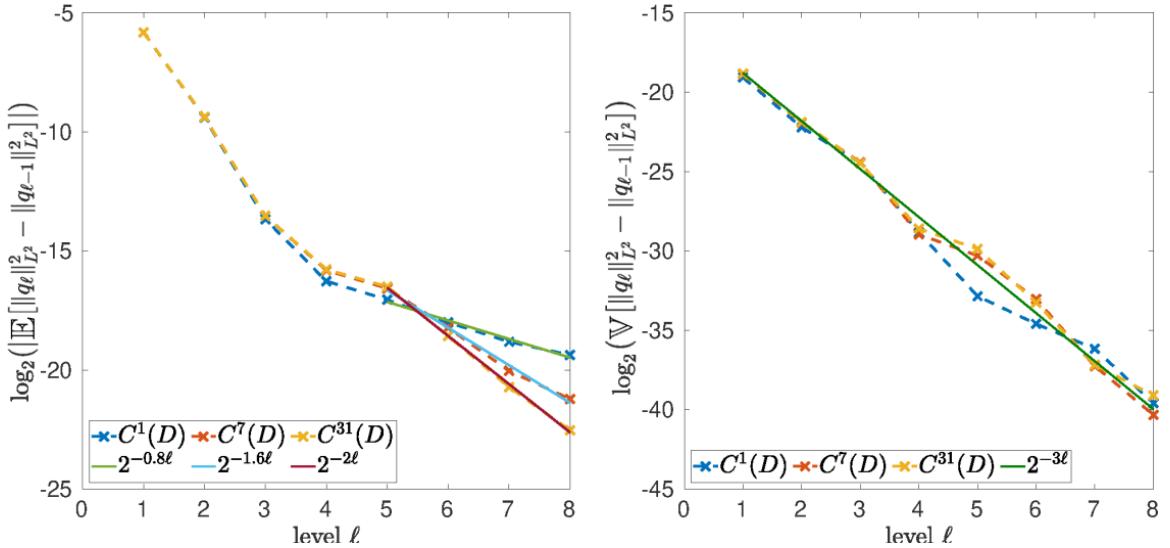


Figure 4.10: Convergence behaviour of the FEM approximation to the solution of (4.65) with  $p$ -refinement in 2D. The approximate FEM solution  $q_\ell$  on level  $\ell$  is obtained by using Lagrange elements of degree  $p = \ell$ . For the polynomial degrees considered we are only able to observe a pre-asymptotic behaviour in which the convergence is geometric. The straight lines (dashed and full) in the left plot indicate the estimated convergence order of the expected value (for  $\nu = 1$  and  $\nu = 31$  respectively). The straight line in the right plot indicates the estimated convergence order of the variance for all the values of  $\nu$  considered.

## 4.6 Chapter conclusions

In this chapter, we have presented a new sampling technique for efficient computation of the action of white noise realisations, even when coupled samples are required within an MLMC framework. This technique applies for general  $L^2$ -conforming finite element spaces, and allows the coupling of samples between non-nested meshes without resorting to a computationally costly interpolation or projection step. The numerical results show that our technique works well in practice: the convergence orders observed agree with existing theory, the number of supermesh cells grows linearly with the finer parent mesh size, the covariance structure of the sampled fields converges to the exact Matérn covariance and the consistency of the telescoping sum is respected. We note that our sampling technique is not limited to Matérn field sampling, but extends naturally to any application in which spatial white noise realisations are needed within a finite element framework.

In the following chapter we extend the method to the multilevel quasi Monte Carlo case.

# Chapter 5

## Multilevel quasi Monte Carlo methods for white noise driven elliptic PDEs

The contents of this chapter are novel unless otherwise mentioned.

All the numerical results in this chapter have been produced with the software FEMLMC, a software for parallel forward uncertainty quantification for PDEs with random coefficients entirely written by the author that contains all the new methods presented in this thesis.

In this chapter we present the extension of the methods presented in Chapter 4 to the multilevel quasi Monte Carlo case (MLQMC). The idea is to express white noise as a Haar wavelet series expansion that we divide in two parts. The first part is sampled using quasi-random points and contains a finite number of terms in order of decaying importance to ensure good QMC convergence. The second part is a correction term which is sampled using standard pseudo-random numbers. We show that the MLQMC method proposed performs better than MLMC in practical experiments.

### 5.1 Related work

In this chapter we consider the solution of elliptic PDEs in which Matérn fields sampled via the SPDE approach appear as coefficients. For instance, a typical problem in the UQ literature reads: find  $\mathbb{E}[P]$ , where  $P(\omega) = \mathcal{P}(p)$  and  $\mathcal{P}$  is a Fréchet differentiable functional of the function  $p$  that satisfies,

$$-\nabla \cdot (F(u(\mathbf{x}, \omega)) \nabla p(\mathbf{x}, \omega)) = f(\mathbf{x}), \quad \mathbf{x} \in G \subset \subset \mathbb{R}^d, \quad \omega \in \Omega, \quad (5.1)$$

with suitable boundary conditions, where  $u(\mathbf{x}, \omega)$  is a Matérn field sampled by solving the Whittle SPDE (2.41) over  $D \supset G$ , the function  $f$  and the domain  $G$  are suitably smooth and  $F \in C^0(\mathbb{R})$  is a positive locally Lipschitz function.

The main focus of this chapter is the presentation of a new QMC and MLQMC method for the estimation of  $\mathbb{E}[P]$  based on the efficient sampling of the white noise

term in (2.41) with a hybrid quasi/pseudo-random sequence. As mentioned in Section 2.1, it is extremely important for QMC applications for the QMC integrand to have low effective dimensionality and to order the QMC integrand variables in order of decaying importance. For this reason, a common approach in the existing literature about MLQMC methods for elliptic PDEs is the expansion of the random field coefficients as an infinite series of basis functions of  $L^2(D)$  that naturally exposes the leading order dimensions in the integrands [47, 75, 108, 109]. If the random field is smooth, the coefficients in the (e.g. Karhunen-Loëve) expansion quickly decay and a truncated expansion provides both the variable ordering and the low-effective dimensionality required by QMC methods.

When using the SPDE approach, the only source of randomness is white noise and we therefore must expand  $\dot{W}$  to achieve the required variable ordering. In this case, the KL-expansion does not provide a feasible route since white noise is not smooth and the eigenvalues in the expansion do not decay. A good alternative in this case is offered by a wavelet expansion of  $\dot{W}$ .

Wavelets in general form a multi-resolution orthogonal basis of  $L^2(D)$  and are commonly employed within QMC algorithms as their hierarchical structure exposes the leading order dimensions in the integrands while allowing fast  $O(m)$  or  $O(m \log m)$  complexity operations (according to whether the wavelets are compactly supported or not [45]). A classical example on the efficacy of wavelet expansions of white noise (in time) within a QMC method is offered by the Lévy-Ciesielski (or Brownian bridge) construction of Brownian motion. Ubiquitous in Mathematical Finance, it is commonly used to solve stochastic differential equations with QMC [67, 71]. Inspired by this technique, we choose to expand white noise into a Haar wavelet expansion<sup>1</sup>, although the generalisation of our approach to higher degree wavelets should be straight-forward.

In a MLQMC framework, wavelets are used by Kuo et al. to sample random fields efficiently, yielding a cost per sample of  $O(m \log m)$  using nested grids [108]. In [88], Hermann and Schwab use a truncated wavelet expansion of white noise to sample Gaussian fields with the SPDE approach within a nested MLQMC hierarchy. Their work is possibly the closest to ours as they also work with the SPDE approach to Matérn field sampling and use a wavelet expansion of white noise [88].

Generally speaking, all the randomised MLQMC methods for elliptic PDEs presented in the above papers are strongly theory-oriented. They use randomly shifted lattice rules and derive MLQMC complexity bounds using a pure QMC approach, truncated expansions and nested hierarchies on simple geometries. Our work is different in spirit and strategy. Firstly, our focus is practice-oriented and we do not derive any MLQMC complexity estimates<sup>2</sup>, but we design our method to work in the general case in which the multilevel hierarchy is non-nested and the domain geometries are non-trivial. Nevertheless, we can still sample white noise (and consequently the Matérn field) in linear cost complexity (or log-linear, according to the Haar wavelet

---

<sup>1</sup>Note that the hat functions used in the Lévy-Ciesielski construction are piecewise linear wavelets, their derivatives are Haar wavelets and white noise in time is the derivative of Brownian motion.

<sup>2</sup>Some of our numerical experiments are outside of the MLQMC convergence theory presented in [88].

type). We leave the choice of QMC point sequence to the user.

Secondly, we handle the expansion differently: we do not just truncate it, but we work with the whole infinite expansion of white noise by adding a correction term to the truncation. The truncation term is finite-dimensional and we sample it with a randomised low-discrepancy sequence; the correction term is infinite-dimensional and a QMC approach is not feasible. However, the covariance of the correction is known and we can sample it using pseudo-random numbers with an extension of the technique presented in Chapter 4.

We therefore adopt a hybrid MC/QMC approach. The advantage of doing so is that we can sample white noise exactly, independently from the truncation level and the wavelet degree considered (e.g. while we use Haar wavelets, Hermann and Schwab in [88] consider higher degree wavelets), without introducing any additional bias into the MLQMC estimate. In contrast, in the aforementioned MLQMC algorithms the expansion must be truncated after enough terms to make the truncation error negligible. Naturally, this advantage comes at a price: since we are using pseudo-random numbers as well, the asymptotic convergence rate of our method with respect to the number of samples  $N$  is still the standard MC rate of  $O(N^{-1/2})$ . Nevertheless, we show that large computational gains can be recovered in practice in a pre-asymptotic QMC-like regime in which the convergence rate is  $O(N^{-\chi})$ ,  $\chi \geq 1/2$ , and we derive a partial convergence result that explains this behaviour in the QMC case.

Wavelets are used in both [108] and [88], but no comment is made about how to work with the wavelet basis in practice if this is not nested within the FEM approximation subspace. In fact, if the field is sampled on a structured grid, this would only typically be piecewise smooth and naively using or transferring such a low-regularity field onto a non-nested grid might cause the quadrature/interpolation/projection error to harm convergence<sup>3</sup>. Furthermore, when working with complex geometries and graded meshes it is desirable for the sampled Matérn field to have the same local accuracy as the solution of the PDE of interest (e.g. (5.1)). For this purpose, it is advantageous to sample the Matérn field on an unstructured mesh with similar properties as the FEM mesh on which it is needed. To address this problem, we again adopt the embedded mesh technique by Osborn et al. [132] so that in the MLQMC hierarchy each mesh of  $G$  is nested within the corresponding mesh of  $D$  and we deal with the non-nestedness of the FEM and wavelet spaces via a supermesh construction. In the independent white noise realisation case we construct a two-way supermesh between the graded FEM mesh of interest and a uniform “wavelet” mesh and we sample white noise in a consistent way between the FEM and the wavelet subspaces<sup>4</sup>. In the MLQMC coupled realisation case, we construct a three-way supermesh between the two non-nested FEM meshes and the “wavelet” mesh.

The supermesh constructions can be simplified when the meshes involved are nested and if all meshes are nested no supermesh is required. In any case, owing to Theorem 4.3, the number of supermesh cells is still linear in the number of cells of the

---

<sup>3</sup>Note that this is not a problem for fields with low regularity since the FEM error dominates in this case, cf [74]

<sup>4</sup>This is done in a similar way as in the coupled realisation case in Chapter 4. This time we replace one of the FEM subspaces with the wavelet subspace.

parent meshes. We remark that the same supermeshing strategy can be employed to sample the truncated white noise expansion used in [88] in the general non-uniform case as our technique easily generalises to higher degree wavelets.

This chapter is structured as follows: in Section 5.2 we introduce the Haar wavelet expansion of white noise and its splitting into a truncated term and a correction term. In Section 5.3 we introduce our sampling technique for independent white noise realisations. A partial QMC convergence result is described in Section 5.4, in which we show that the QMC integration error also splits in two terms, one QMC-like term converging at a faster-than-MC rate and one MC-like term that decays at a standard  $O(N^{-1/2})$  rate with respect to the number of samples and exponentially fast with respect to the wavelet series truncation. We extend the white noise sampling method to MLQMC in Section 5.5, where we show how coupled white noise realisations can be sampled efficiently. The algorithms introduced are supported by numerical results, which we present and discuss in Section 5.6. We conclude the chapter with a brief summary of the methods and results presented in Section 5.7.

## 5.2 Haar wavelet expansion of spatial white noise

Low-discrepancy sequences are extremely uniform in the first few dimensions and in low-dimensional projections, but less so across the whole hypercube. Consequently QMC works best when the integrand has low effective dimension [98]. For good QMC convergence we need to order the dimensions of the QMC integrand in order of decaying importance so that the largest error components are on the first dimensions [47, 71]. In what follows we expand white noise into a Haar wavelet series so that the hierarchical structure of Haar wavelets can naturally provide the variable ordering needed for QMC integration.

We start by introducing the Haar wavelet basis. Let  $\mathbb{1}_A(x)$ , be the indicator function of a set  $A$  and let  $\Psi(x)$  for  $x \in \mathbb{R}$  be the Haar mother wavelet (see Figure 5.1),

$$\Psi(x) = \mathbb{1}_{[0,1/2)}(x) - \mathbb{1}_{[1/2,1)}(x) = \begin{cases} 1, & 0 \leq x < 1/2, \\ -1, & 1/2 \leq x < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (5.2)$$

Let  $\bar{\mathbb{N}} = \{-1\} \cup \mathbb{N}$  and let  $x^+ = \max(x, 0)$ . The Haar wavelets  $H_{l,n}$  for  $l \in \bar{\mathbb{N}}$ ,  $n = 0, \dots, (2^l - 1)^+$  can be expressed in terms of the mother wavelet through shifting and rescaling as follows.

$$\begin{cases} H_{-1,0}(x) = \mathbb{1}_{[0,1)}(x), & l = -1, n = 0 \\ H_{l,n}(x) = 2^{l/2}\Psi(2^lx - n), & l \in \mathbb{N}, n = 0, \dots, 2^l - 1. \end{cases} \quad (5.3)$$

The Haar wavelets have support size  $|\text{supp}(H_{l,n})| = 2^{-l^+}$  and form an orthonormal basis of  $L^2((0, 1))$ . The Haar system can be generalised to higher dimensions by

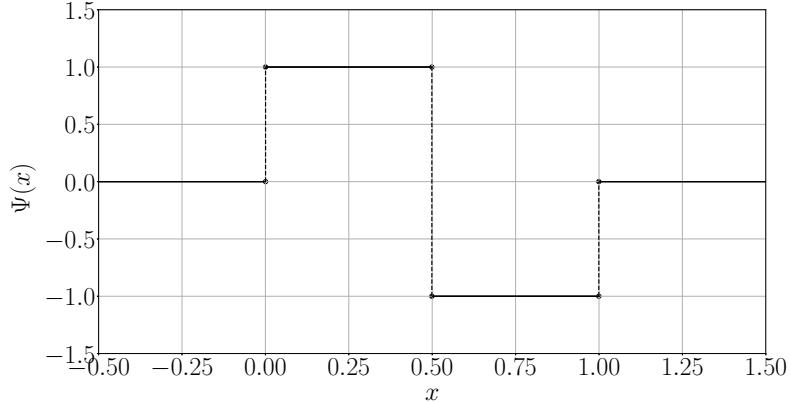


Figure 5.1: *Haar mother wavelet.*

taking the tensor product of the 1D Haar basis with itself: let  $\mathbf{l} \in \bar{\mathbb{N}}^d$  and  $\mathbf{n} \in \mathbb{N}$  we can define the family of  $d$ -dimensional Haar wavelets  $H_{\mathbf{l},\mathbf{n}}(\mathbf{x})$  for  $\mathbf{x} \in \mathbb{R}^d$  as

$$H_{\mathbf{l},\mathbf{n}}(\mathbf{x}) = \bigotimes_{i=1}^d H_{l_i, n_i}(x_i), \quad \text{with } n_i \in \{0, \dots, (2^{l_i} - 1)^+\} \forall i. \quad (5.4)$$

The  $d$ -dimensional Haar wavelets have support size  $|\text{supp}(H_{\mathbf{l},\mathbf{n}})| = \prod_{i=1}^d 2^{-l_i^+} = 2^{-|\mathbf{l}^+|_1}$  and they form an orthonormal basis of  $L^2((0, 1)^d)$ .

*Remark 5.1* (General boxed domains). It is also possible to construct an orthonormal basis for  $L^2((0, a))$  by defining the wavelets as

$$\begin{cases} H_{-1, -1}^a(x) = \frac{1}{\sqrt{a}} \mathbb{1}_{[0, a]}(x), & l = -1, n = 0 \\ H_{l, n}^a(x) = \frac{2^{l/2}}{\sqrt{a}} \Psi\left(\frac{2^l}{a}x - n\right), & l \in \mathbb{N}, n = 0, \dots, 2^l - 1, \end{cases} \quad (5.5)$$

where  $x \in [0, a]$ . In other words, it is sufficient to rescale:  $H_{l, n}^a = a^{-1/2} H_{l, n}(x/a)$ .

*Remark 5.2* (On the compactly supported Haar wavelet basis). It is worth mentioning that there is another way of extending the 1D Haar basis to higher dimensions. The  $d$ -dimensional wavelets we have introduced are sometimes called the *standard* Haar basis and are not compactly supported for  $d > 1$ . This can be disadvantageous from a numerical point of view since there are exactly  $(\mathcal{L} + 2)^d$  non-zero wavelets of integer level  $|\mathbf{l}|_\infty \leq \mathcal{L}$  at any point  $x \in D$ . If the evaluation of the wavelets becomes an issue, a better option is to use the *non-standard* (or compactly supported) Haar basis, first introduced by Daubechies in [45]. The non-standard Haar wavelets are obtained by direct scaling and shifting of the  $d$ -dimensional standard Haar wavelets of level 0 and have the advantageous property of being compactly supported for all  $\mathcal{L}$  (for further details, see [23, 45]). Only  $(2^d - 1)\mathcal{L} + 2^d$  non-standard wavelets are non-zero at any given point  $x \in D$ . For a numerical comparison between the two types of bases, see e.g. [23]. Although in our numerical implementation we used the standard basis, the algorithms we are about to introduce work for both.

Let  $|\mathbf{l}| = \max_i(l_i)$ . We can now express white noise in  $[0, 1]^d$  as the generalised stochastic field of type II (cf. Definition 2.5) given by the Haar wavelet expansion,

$$\dot{W} = \sum_{|\mathbf{l}|=-1}^{\infty} \sum_{\mathbf{n}=0}^{(2^l-1)^+} z_{\mathbf{l},\mathbf{n}}(\omega) H_{\mathbf{l},\mathbf{n}}(\mathbf{x}), \quad (5.6)$$

where  $z_{\mathbf{l},\mathbf{n}} = \langle \dot{W}, H_{\mathbf{l},\mathbf{n}} \rangle$  for all  $\mathbf{l}, \mathbf{n}$ , making the  $z_{\mathbf{l},\mathbf{n}}$  be i.i.d. standard Gaussian random variables. The second summation in (5.6) is to be intended as the sum over all  $\mathbf{n}$  with components  $n_i$  such that  $0 \leq n_i \leq \max(2^{l_i} - 1, 0)$  for all  $i$ . The same Haar wavelet expansion can be used to write the action of white noise as (recall the definition of type I generalised stochastic fields, Definition 2.4)

$$\langle \dot{W}, v \rangle = \sum_{|\mathbf{l}|=-1}^{\infty} \sum_{\mathbf{n}=0}^{(2^l-1)^+} z_{\mathbf{l},\mathbf{n}}(H_{\mathbf{l},\mathbf{n}}, v), \quad \forall v \in L^2(D). \quad (5.7)$$

Let  $\mathcal{L} \in \bar{\mathbb{N}}$ . We now divide the series in two terms,

$$\dot{W} = \dot{W}_{\mathcal{L}} + \dot{W}_R = \sum_{|\mathbf{l}|=-1}^{\mathcal{L}} \sum_{\mathbf{n}=0}^{(2^l-1)^+} z_{\mathbf{l},\mathbf{n}}(\omega) H_{\mathbf{l},\mathbf{n}}(\mathbf{x}) + \sum_{|\mathbf{l}|=\mathcal{L}+1}^{\infty} \sum_{\mathbf{n}=0}^{(2^l-1)^+} z_{\mathbf{l},\mathbf{n}}(\omega) H_{\mathbf{l},\mathbf{n}}(\mathbf{x}). \quad (5.8)$$

Here the term  $\dot{W}_{\mathcal{L}}$  is a “proper” stochastic field in  $L^2(\Omega, L^2(D))$ , while  $\dot{W}_R$  is again a generalised stochastic field of type I with action defined by

$$\langle \dot{W}_R, v \rangle = \sum_{|\mathbf{l}|=\mathcal{L}+1}^{\infty} \sum_{\mathbf{n}=0}^{(2^l-1)^+} z_{\mathbf{l},\mathbf{n}}(H_{\mathbf{l},\mathbf{n}}, v), \quad \forall v \in L^2(D). \quad (5.9)$$

The idea is then to sample the Gaussian variables in the expression for  $\dot{W}_{\mathcal{L}}$  by using a hybrid QMC/MC combination of quasi-random (e.g. Sobol) and pseudo-random numbers and to sample  $\dot{W}_R$  with pseudo-random numbers only by extending the work in Chapter 4.

The reasoning behind this splitting is that it is important to keep the dimensionality of the QMC point sequence relatively low: 1) as we will see in the next section, the sampling of  $\dot{W}$  expressed this way requires a supermesh construction and smaller dimensions imply faster  $\dot{W}$  samples; 2) some QMC point sequences cannot readily be sampled in high dimensions<sup>5</sup> and 3) the approximation properties of some quasi-random sequences deteriorate as the dimensionality grows [47, 71].

### 5.3 Sampling independent realisations for QMC

Again, we want to solve equation (2.41), namely

$$\begin{aligned} Lu &= u - \kappa^{-2} \Delta u = \eta \dot{W}, \quad \mathbf{x} \in D, \\ u &= 0, \quad \mathbf{x} \in \partial D. \end{aligned} \quad (5.10)$$

---

<sup>5</sup>For example, the state-of-the-art Sobol’ sequence generator, Broda, can generate the largest dimensional Sobol’ sequences with 65536 dimensions [160]. This might still be too low for an infinite-dimensional PDE setting.

From now on we introduce the simplifying assumptions that  $D = [0, 1]^d$ . Relaxing this assumption to general boxed domains is straight-forward, but considering more general cases is non-trivial (more on this later on in Remark 5.19). Since we are interested in Matérn field sampling we are free to choose any domain shape (as we are truncating  $\mathbb{R}^d$  to an arbitrary  $D$  anyway) so this is not really a restriction. It is useful for what comes next to introduce the concept of a Haar mesh (see Figure 5.2):

**Definition 5.1** (Haar mesh). Let  $D = [0, 1]^d$  and let  $\mathcal{L} \in \bar{\mathbb{N}}$ . The Haar mesh  $D_{\mathcal{L}}$  is the uniform quadrilateral mesh of  $D$  whose cells are all regular polyhedra of volume  $|\square_H| = 2^{-d(\mathcal{L}+1)}$ . Note that for a given  $\mathcal{L}$  there are exactly as many cells in  $D_{\mathcal{L}}$  as terms in the wavelet expansion (5.8) for  $\dot{W}_{\mathcal{L}}$ , namely  $\mathcal{N}_{\mathcal{L}} = 2^{d(\mathcal{L}+1)}$ .

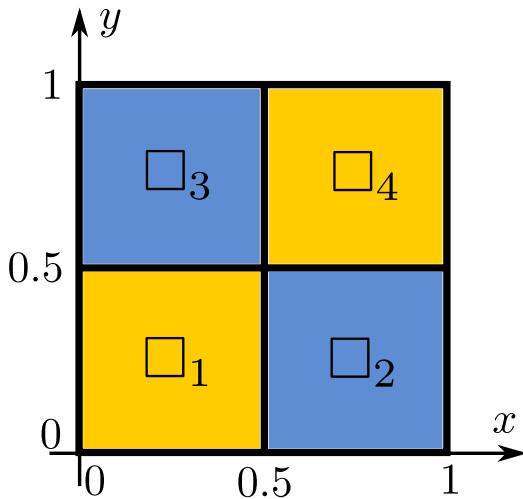


Figure 5.2: The Haar mesh in the  $d = 2$ ,  $\mathcal{L} = 0$  case. The Haar cells are coloured according to the values of the  $H_{0,0} = \Psi(x)\Psi(y)$  wavelet: yellow for  $+1$ , blue for  $-1$ .

Let  $D_h$  be a mesh of  $D$ , not necessarily nested within the Haar mesh  $D_{\mathcal{L}}$ , let  $V \subseteq L^2(D)$  and let  $V_h = \text{span}(\phi_1, \dots, \phi_m) \subseteq V$  be the FEM subspace used to solve equation (5.10) on  $D_h$ . In what follows we will refer to  $D_h$  as the *FEM mesh* and we assume for simplicity that there are always  $m_e$  degrees of freedom of  $V_h$  on each cell of  $D_h$ .

We can now focus on the sampling of  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$ . As in our previous work (see Chapter 4), we address the sampling of the action of  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$  against the basis functions of  $V_h$ , i.e. we sample the two zero-mean Gaussian vectors  $\mathbf{b}_{\mathcal{L}}$  and  $\mathbf{b}_R$  with entries given by

$$(\mathbf{b}_{\mathcal{L}})_i = (\dot{W}_{\mathcal{L}}, \phi_i), \quad (\mathbf{b}_R)_i = \langle \dot{W}_R, \phi_i \rangle. \quad (5.11)$$

Note the difference between the brackets. This is because the sum in the expression (5.8) for  $\dot{W}_{\mathcal{L}}$  is finite, hence  $\dot{W}_{\mathcal{L}}$  is a ‘proper’ Gaussian field in  $L^2(\Omega, L^2(D))$ , while  $\dot{W}_R$  is still a generalised stochastic field (of both types) and the angle bracket notation (cf. Definition 2.4) is still needed.

*Remark 5.3.* When the FEM mesh is non-nested within the Haar mesh, the computation of integrals involving products of Haar wavelets and FEM basis functions becomes difficult since the quantities involved are only piecewise smooth. To solve this

problem, a supermesh construction between the FEM and Haar meshes is required to split the support of the wavelets and FEM basis functions into smooth portions that can be integrated with high accuracy. For this reason, the methods we present in this chapter always rely on a supermesh construction whenever the two meshes are not nested. From now on we assume without loss of generality that the support of each  $\phi_i \in V_h$  is entirely contained in a single Haar mesh cell, i.e. that each basis function  $\phi_i \in V_h$  has been split into the sum of the restrictions of  $\phi_i$  to each cell of  $D_{\mathcal{L}}$  by using a supermesh construction. We will indicate with  $S_h$  the constructed supermesh and with abuse of notation we will still assume these “splitted” basis functions to be  $m$ .

*Remark 5.4.* Note that the supermesh construction is not only specific to our algorithms, but it might also be required by other methods that produce Gaussian field samples on a uniform grid to transfer (or integrate) the sampled field to (or over) the non-uniform mesh of interest. If the sampled fields have lower than  $H^1(D)$  regularity non-nested interpolation can also be used without harming the overall convergence rate [74], although this operation is likely to marginally increase the bias.

### 5.3.1 Sampling of $\dot{W}_{\mathcal{L}}$

We now consider the efficient sampling of  $\dot{W}_{\mathcal{L}}$ . Let  $V_H = \text{span}(\psi_1, \dots, \psi_{N_{\mathcal{L}}})$  with  $\psi_k = \mathbf{1}_{\square_k}$  be the space of piecewise constant functions over the cells  $\square_k$  of the Haar mesh  $D_{\mathcal{L}}$ . It turns out that  $\dot{W}_{\mathcal{L}} \in V_H$  almost surely and that therefore it can be expressed in terms of the basis functions of  $V_H$  as  $\dot{W}_{\mathcal{L}} = \sum_{k=1}^{N_{\mathcal{L}}} w_k \psi_k$ , where  $w_k$  is the value of  $\dot{W}_{\mathcal{L}}$  over the Haar cell  $\square_k$ . In practice, rather than computing the inner products of each Haar wavelet with the basis functions of  $V_h$ , it is more straightforward to just compute each entry of  $\mathbf{b}_{\mathcal{L}}$  as  $(\mathbf{b}_{\mathcal{L}})_i = w_{\kappa(i)} \int_D \phi_i \, d\mathbf{x}$ . Here  $\kappa(i)$  is the index  $k$  of the Haar cell that contains the support of  $\phi_i$  and  $w_{\kappa(i)}$  must be computed from each sample of the coefficients in the expansion for  $\dot{W}_{\mathcal{L}}$ . Before explaining how this is actually done in practice, we prove that  $\dot{W}_{\mathcal{L}}$  can be interpreted as the projection of white noise onto  $V_H$  and therefore  $\dot{W}_{\mathcal{L}}$  does indeed belong to  $V_H$ .

**Lemma 5.1.** *Let  $V_H = \text{span}(\psi_1, \dots, \psi_{N_{\mathcal{L}}})$  with  $\psi_i = \mathbf{1}_{\square_i}$  be the space of piecewise constant functions over the cells  $\square_i$  of the Haar mesh  $D_{\mathcal{L}}$ . Let  $P_H$  be the  $L^2$  projection onto  $V_H$  and define the projected white noise  $P_H \dot{W}$  (similarly as in Lemma 3.4) as follows (recall that the angle bracket notation indicates the action of the white noise operator),*

$$(P_H \dot{W}, v) := \dot{W}(P_H v) \equiv \langle \dot{W}, P_H v \rangle, \quad \forall v \in L^2(D). \quad (5.12)$$

*We then have that  $\dot{W}_{\mathcal{L}} \equiv P_H \dot{W}$  in  $L^2(\Omega, L^2(D))$ .*

*Proof.* We note that all the Haar wavelets in the expansion for  $\dot{W}_{\mathcal{L}}$  can be represented as a linear combination of basis functions of  $V_H$ . Since there are exactly as many wavelets as basis functions of  $V_H$  (see Definition 5.1) and since these wavelets are linearly independent, we conclude that the Haar wavelets form a basis of  $V_H$ .

Therefore  $\dot{W}_{\mathcal{L}} \in L^2(\Omega, V_H)$  and  $\langle \dot{W}_R, v \rangle = 0$  for all  $v \in V_H$  (cf. equation (5.9)). Furthermore, for all  $v \in L^2(D)$ ,

$$(\dot{W}_{\mathcal{L}}, v) = (\dot{W}_{\mathcal{L}}, P_H v + v^\perp) = (\dot{W}_{\mathcal{L}}, P_H v) = \langle \dot{W}, P_H v \rangle =: (P_H \dot{W}, v), \quad (5.13)$$

almost surely since  $\langle \dot{W}_R, P_H v \rangle = 0$  for all  $v \in L^2(D)$  (cf. equation (5.9)). Here we used the fact that all  $v \in L^2(D)$  can be split as  $v = P_H v + v^\perp$ , where  $v^\perp \in V_H^\perp$ .  $\square$

Note that the dimension of the space  $V_H$  is  $\mathcal{N}_{\mathcal{L}} = 2^{d(\mathcal{L}+1)}$ , which is quite a large number. If we were to sample  $\dot{W}_{\mathcal{L}}$  with a pure QMC approach, we would therefore need a  $\mathcal{N}_{\mathcal{L}}$ -dimensional QMC point sequence, that we might not be able to sample given the restrictions of some modern QMC point generators (cf. Remark 5.6). In the interest of reducing the QMC dimension, we reorder the terms in the expression for  $\dot{W}_{\mathcal{L}}$  in (5.8) with respect to a total degree ordering rather than a full tensor grid ordering (i.e. we reorder them with respect to the 1 norm of  $\mathbf{l}$  rather than the max norm) and we then only sample the Haar coefficients with  $|\mathbf{l}|_1 \leq \mathcal{L}$  which are always much less than those with  $|\mathbf{l}| \leq \mathcal{L}$ , cf. Remark 5.6. To fix the notation, for  $s \in \bar{\mathbb{N}}$ , we define the set

$$\mathcal{H}(s) := \{\mathbf{l} \in \bar{\mathbb{N}}^d : |\mathbf{l}|_1 = s, |\mathbf{l}| \leq \mathcal{L}\}, \quad \text{with } \mathcal{H}(-1) = \{-\mathbf{1}\}. \quad (5.14)$$

This is the set of all Haar level vectors in the expansion for  $\dot{W}_{\mathcal{L}}$  of a given total degree. After reordering the terms in the expression for  $\dot{W}_{\mathcal{L}}$  in (5.8) with respect to the 1-norm of  $\mathbf{l}$  we obtain:

$$\dot{W}_{\mathcal{L}} = \sum_{s=0}^{s=\mathcal{L}} \sum_{\mathbf{l} \in \mathcal{H}(s)} \sum_{\mathbf{n}=0}^{(2^{\mathbf{l}}-1)^+} z_{\mathbf{l},\mathbf{n}}(\omega) H_{\mathbf{l},\mathbf{n}}(\mathbf{x}) + \sum_{s=\mathcal{L}+1}^{s=d\mathcal{L}} \sum_{\mathbf{l} \in \mathcal{H}(s)} \sum_{\mathbf{n}=0}^{(2^{\mathbf{l}}-1)^+} z_{\mathbf{l},\mathbf{n}}(\omega) H_{\mathbf{l},\mathbf{n}}(\mathbf{x}), \quad (5.15)$$

where we sample the coefficients in the first sum on the right hand side with QMC and the remaining coefficients with a standard MC approach. We therefore adopt a hybrid sampling technique for  $\dot{W}_{\mathcal{L}}$ .

In order to achieve good convergence with respect to the number of QMC samples we order the QMC dimensions according to  $s$  so that the first dimension corresponds to  $z_{-\mathbf{1},0}$ , the second batch of dimensions correspond to the terms with  $s = 0$ , the third batch of dimensions to the terms with  $s = 1$  and so on up until  $s = \mathcal{L}$ . We map each sampled low-discrepancy sequence point (in our case Sobol with digital shifting [71]) to a Gaussian-distributed sequence by applying the inverse Normal CDF. We sample the remaining coefficients independently using a pseudo-random number generator. To fix ideas, we show a schematic of our sampling choices for the coefficients of  $\dot{W}$  in Figure 5.3.

*Remark 5.5.* Note that both orderings (max and 1 norm) are essential: the white noise expansion (5.6) must be split according to the max norm so that  $\dot{W}_{\mathcal{L}}$  can be interpreted as the projection of white noise onto  $V_H$ . In principle, the max norm could also be used to enforce the ordering required for QMC convergence. However, this would involve sampling for an extremely high dimensional QMC point sequence.

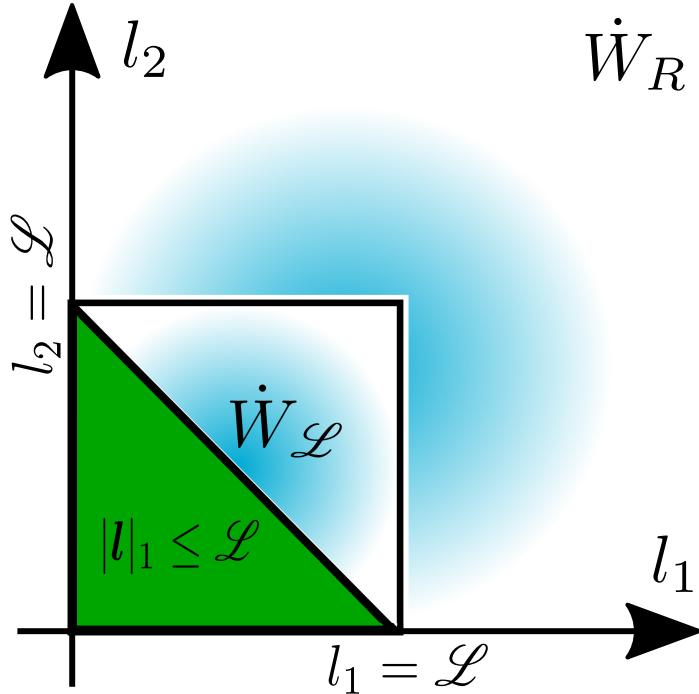


Figure 5.3: A schematic of the sampling strategy for the Haar coefficients of  $\dot{W}$  in 2D. The coefficients in the square are the coefficients for  $\dot{W}_{\mathcal{L}}$ , while those in the unbounded “L-shaped” domain belong to  $\dot{W}_R$ . The region coloured in green corresponds to the coefficients with  $|l|_1 \leq \mathcal{L}$  which are sampled with a low-discrepancy sequence. The others, corresponding to the regions in light blue, are sampled with independent pseudo-random numbers.

Ordering the terms in the expansion for  $\dot{W}_{\mathcal{L}}$  according to the 1 norm instead allows us to still enforce a good QMC ordering while reducing the dimension of the QMC point rule used.

*Remark 5.6.* To get an idea of the numbers, note that there are  $2^{d(\mathcal{L}+1)}$  wavelets satisfying  $|l| \leq \mathcal{L}$ , but only  $2^{\mathcal{L}-1}(\mathcal{L} + 3)$  and  $2^{\mathcal{L}-2}(\mathcal{L}^2 + 9\mathcal{L} + 16)$  satisfying  $|l|_1 \leq \mathcal{L}$  in 2D and 3D respectively<sup>6</sup>. The Sobol’ generator we use is based on Joe and Kuo’s direction number sets [98] and can get up to 21201 dimensions. If we wanted to sample all the coefficients in  $\dot{W}_{\mathcal{L}}$  with Sobol’ points, we would be unable to go past  $\mathcal{L} = 6$  in 2D and  $\mathcal{L} = 3$  in 3D. By sampling only the coefficients with  $|l|_1 \leq \mathcal{L}$  we can get up to  $\mathcal{L} = 11$  in 2D and  $\mathcal{L} = 9$  in 3D. This argument is still valid even if we use the more advanced generator Broda that can get up to 65536 dimensions [160].

---

<sup>6</sup>We omit the derivation of these expressions as it is a simple, yet tedious counting exercise. Note that in 1D  $|\cdot| \equiv |\cdot|_1$ .

We now propose the following algorithm for sampling  $\dot{W}_{\mathcal{L}}$ :

**Algorithm for the sampling of  $\dot{W}_{\mathcal{L}}$ :**

1. Compute the supermesh between the FEM mesh and the Haar mesh and split the support of the basis functions of  $V_h$  to obtain  $\{\phi_i\}_{i=1}^m$  each of which with support entirely contained within a single Haar cell. Compute the scalar map  $\kappa(i)$  that maps each  $i$  to the index  $k$  of the Haar cell  $\square_k$  that contains the support of  $\phi_i$  and compute  $\int_D \phi_i \, d\mathbf{x}$  for all  $i = 1, \dots, m$ . This step can be done once and for all as an offline step.
2. Sample the vector  $\mathbf{z}_{\mathcal{L}} \in \mathbb{R}^{\mathcal{N}_{\mathcal{L}}}$  of the coefficients in the expression (5.8) for  $\dot{W}_{\mathcal{L}}$  as  $\mathbf{z}_{\mathcal{L}} = [\mathbf{z}_{\text{QMC}}^T, \mathbf{z}_{\text{MC}}^T]^T$ , where  $\mathbf{z}_{\text{QMC}}$  is obtained by applying the inverse Gaussian CDF to a randomised low-discrepancy sequence point of dimension equal to the number of coefficients with  $|\mathbf{l}|_1 \leq \mathcal{L}$  and  $\mathbf{z}_{\text{MC}}$  is independently sampled with a pseudo-random number generator.
3. We sample the values  $w_k$  of  $\dot{W}_{\mathcal{L}}$  over each Haar mesh cell  $\square_k$  as follows. Let  $J(\mathbf{l}, \mathbf{n})$  be the index map that given  $(\mathbf{l}, \mathbf{n})$  returns the index  $j$  such that  $z_{l,n} = (\mathbf{z}_{\mathcal{L}})_j$  (the two vectors are the same up to reordering) and define  $\mathbf{m}_k \in \mathbb{R}^d$  to be the coordinate vector of the midpoint of  $\square_k$ . For each  $k = 1, \dots, \mathcal{N}_{\mathcal{L}}$  and  $\mathbf{l}$  with  $|\mathbf{l}| \leq \mathcal{L}$ , there is only one wavelet with level vector  $\mathbf{l}$  with non-zero support over  $\square_k$ . For  $i = 1, \dots, d$ , its wavelet number is given by  $(\bar{\mathbf{n}}_k(\mathbf{l}))_i = \lfloor (\mathbf{m}_k)_i 2^{l_i} \rfloor$  and its sign over  $\square_k$  by  $\bar{s}_k(\mathbf{l}) = \prod_{i=1}^d s_k(\mathbf{l}_i)$ , where the  $s_k(\mathbf{l}_i)$  are the signs of the 1D Haar wavelets in the tensor product for  $H_{\mathbf{l}, \bar{\mathbf{n}}_k(\mathbf{l})}$ , namely

$$s_k(\mathbf{l}_i) = 1 - 2(\lfloor (\mathbf{m}_k)_i 2^{l_i+1} \rfloor \pmod{2}). \quad (5.16)$$

This expression comes from the fact that Haar wavelets are positive on even Haar cells and negative on odd cells. We set for all  $k = 1, \dots, \mathcal{N}_{\mathcal{L}}$ ,

$$w_k = \sum_{\mathbf{l}=0}^{|\mathbf{l}| \leq \mathcal{L}} \bar{s}_k(\mathbf{l}) \mathbf{z}_{J(\mathbf{l}, \bar{\mathbf{n}}_k(\mathbf{l}))} 2^{|\mathbf{l}|_1/2} \quad (5.17)$$

4. For all  $i = 1, \dots, m$ , set  $(\mathbf{b}_{\mathcal{L}})_i = w_{\kappa(i)} \int_D \phi_i \, d\mathbf{x}$ .

*Remark 5.7* (Exploiting the structure in  $D_{\mathcal{L}}$ ). In point 1 and 3 above we exploit the fact that the Haar mesh is uniform and structured. For instance, we can readily obtain the Haar mesh cell in which any point  $\mathbf{p} \in D$  lies: it belongs to the  $\lfloor (\mathbf{p})_i 2^{\mathcal{L}+1} \rfloor$ -th Haar cell from the origin in the  $i$ -th coordinate direction. When supermeshing this makes the search for candidate intersections [61] inexpensive as we always know for a given cell of  $D_h$  exactly which Haar cells it intersects. The expressions for  $\bar{\mathbf{n}}_k(\mathbf{l})$  and  $\bar{s}_k(\mathbf{l})$  in point 3 above also derive from the same considerations.

*Remark 5.8* (Complexity of the sampling of  $\dot{W}_{\mathcal{L}}$ ). Let  $\mathcal{N}_S$  be the number of cells of the supermesh between the Haar and FEM meshes and let  $m_e$  be the number of degrees of freedom of the FEM basis on each supermesh cell  $e$ . Let  $\mathcal{N}_{\mathcal{L}} = 2^{d(\mathcal{L}+1)}$

be the number of cells in the Haar mesh and let  $N_{\mathcal{L}} = (\mathcal{L} + 2)^d$  be the number of wavelets that are non-zero over a given Haar cell. In general, it is possible to sample  $\dot{W}_{\mathcal{L}}$  in  $O(N_S m_e + N_{\mathcal{L}} \mathcal{N}_{\mathcal{L}})$ . Owing to Theorem 4.3, when  $D_h$  is quasi-uniform we have  $N_S \leq a \mathcal{N}_h + b \mathcal{N}_{\mathcal{L}}$ , where  $\mathcal{N}_h$  is the number of cells of  $D_h$  and  $a, b > 0$ . This gives a cost complexity of order  $O((a \mathcal{N}_h + b \mathcal{N}_{\mathcal{L}}) m_e + N_{\mathcal{L}} \mathcal{N}_{\mathcal{L}})$  which is linear in the number of cells of  $D_h$  and log-linear in the number of cells of  $D_{\mathcal{L}}$  since  $N_{\mathcal{L}} = O((\log_2(\mathcal{N}_{\mathcal{L}})/d)^d)$ . The log-linear term can be replaced with a linear term in case the compactly supported Haar wavelets (cf. Remark 5.2) are used: in this case it is possible to use a multi-dimensional generalisation of the Brownian bridge construction (of which  $\dot{W}_{\mathcal{L}}$  is the derivative) which is well known in the computational finance literature [71].

### 5.3.2 Sampling of $\dot{W}_R$

We now consider the efficient sampling of  $\dot{W}_R$ . Dealing with an infinite summation is complicated. However, we can circumvent this problem by noting that the covariance of  $\dot{W}_R$  is known<sup>7</sup> since, as  $\dot{W}_{\mathcal{L}}$  is independent from  $\dot{W}_R$  by construction, for all  $u, v \in L^2(D)$  we have

$$\mathbb{E}[\langle \dot{W}_R, u \rangle \langle \dot{W}_R, v \rangle] = \mathbb{E}[\langle \dot{W}, u \rangle \langle \dot{W}, v \rangle] - \mathbb{E}[\langle \dot{W}_{\mathcal{L}}, u \rangle \langle \dot{W}_{\mathcal{L}}, v \rangle], \quad (5.18)$$

where the covariance of  $\dot{W}$  is known by Definition 2.6 and the covariance of  $\dot{W}_{\mathcal{L}}$  is given by the following lemma.

**Lemma 5.2.** *Let  $\square_i$  for  $i = 1, \dots, \mathcal{N}_{\mathcal{L}}$  be the  $i$ -th cell of  $D_{\mathcal{L}}$  of volume  $|\square_i| = 2^{d(\mathcal{L}+1)} = |\square_H|$  for all  $i$  (see Definition 5.1). Then, for all  $u, v \in L^2(D)$ ,*

$$\mathcal{C}_{\mathcal{L}}(u, v) = \mathbb{E}[(\dot{W}_{\mathcal{L}}, u)(\dot{W}_{\mathcal{L}}, v)] = \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} \frac{1}{|\square_i|} \int_{\square_i} u \, d\mathbf{x} \int_{\square_i} v \, d\mathbf{x}. \quad (5.19)$$

*Proof.* Let  $P_H$  be the  $L^2$  projection onto  $V_H$ , then for all  $u \in L^2(D)$  we have that  $P_H u = \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} u_i \psi_i$  satisfies

$$(P_H u, v_H) = (u, v_H), \quad \forall v_H \in V_H. \quad (5.20)$$

A standard FEM calculation gives that the coefficients  $u_i$  are given by

$$u_i = \frac{1}{|\square_i|} (u, \psi_i) = \frac{1}{|\square_i|} \int_{\square_i} u \, d\mathbf{x}. \quad (5.21)$$

We conclude by using Lemma 5.1 to show that, for all  $u, v \in V$  such that  $P_H u = \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} u_i \psi_i$  and  $P_H v = \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} v_i \psi_i$ ,

$$\begin{aligned} \mathbb{E}[(\dot{W}_{\mathcal{L}}, u)(\dot{W}_{\mathcal{L}}, v)] &= \mathbb{E}[\langle \dot{W}, P_H u \rangle \langle \dot{W}, P_H v \rangle] = (P_H u, P_H v) \\ &= \sum_{i,j=1}^{\mathcal{N}_{\mathcal{L}}} u_i v_j (\psi_i, \psi_j) = \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} |\square_i| u_i v_i = \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} \frac{1}{|\square_i|} \int_{\square_i} u \, d\mathbf{x} \int_{\square_i} v \, d\mathbf{x}. \end{aligned} \quad (5.22)$$

□

---

<sup>7</sup>As we saw in Chapter 4, knowing the covariance matrix of a Gaussian vector is a key ingredient needed for its sampling.

*Remark 5.9.* The sampling strategies for  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$  presented in this work are conceptually different. In the  $\dot{W}_{\mathcal{L}}$  case we use the Haar wavelet representation to make sure that the variables in the quasi-random sequence are ordered correctly. Therefore the use of the Haar representation is crucial in the sampling of  $\dot{W}_{\mathcal{L}}$ . In the  $\dot{W}_R$  case, instead, the ordering is irrelevant as  $\dot{W}_R$  is sampled by using pseudo-random numbers. For this reason we can “forget” about the wavelet representation in this case and sample  $\dot{W}_R$  as it is done for any standard Gaussian field, i.e. by factorising its covariance matrix after discretisation.

It is then readily shown from Lemma 5.2 and from (5.18) that the covariance of  $\dot{W}_R$  is given by:

$$\mathcal{C}_R(u, v) = \mathbb{E}[\langle \dot{W}_R, u \rangle \langle \dot{W}_R, v \rangle] = (u, v) - \sum_{i=1}^{\mathcal{N}_{\mathcal{L}}} \frac{1}{|\square_i|} \int_{\square_i} u \, d\mathbf{x} \int_{\square_i} v \, d\mathbf{x}, \quad (5.23)$$

for all  $u, v \in L^2(D)$ . From Lemma 5.2 and from Definition 2.6 we deduce that if the supports of  $u$  and  $v$  never share the same Haar mesh cell, then

$$\mathbb{E}[(\dot{W}_{\mathcal{L}}, u)(\dot{W}_{\mathcal{L}}, v)] = \mathbb{E}[\langle \dot{W}, u \rangle \langle \dot{W}, v \rangle] = \mathbb{E}[\langle \dot{W}_R, u \rangle \langle \dot{W}_R, v \rangle] = 0, \quad (5.24)$$

i.e. the action of  $\dot{W}_{\mathcal{L}}$  is exactly the same as the action of white noise in this case and the correction term  $\dot{W}_R$  is not needed. This means that the restrictions of  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$  to separate Haar mesh cells are statistically independent from each other. We can draw a quick comparison with white noise: the restrictions of white noise to disjoint regions of  $D$  are always independent, while the restrictions of  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$  to disjoint regions are still dependent if these regions belong to the same Haar mesh cell. Thanks to this property, we can consider each Haar cell separately and only account for the correlations among the pairings of  $\dot{W}_R$  with test functions that belong to the same cell. Since the computations on separate Haar cells are independent, these operations can be performed simultaneously in parallel.

Before proceeding, we show that  $\mathcal{C}_R$  is a proper covariance function, i.e. that it is positive semi-definite.

**Lemma 5.3.** *The covariance of  $\dot{W}_R$ ,  $\mathcal{C}_R$ , is positive semi-definite.*

*Proof.* With the same notation as in the proof of Lemma 5.2, we have that, for all  $u \in L^2(D)$ ,

$$\mathcal{C}_R(u, u) = \mathbb{E}[(\langle \dot{W}_R, u \rangle)^2] = \mathbb{E}[(\langle \dot{W} - P_H \dot{W}, u \rangle)^2] = \mathbb{E}[(\langle \dot{W}, u - P_H u \rangle)^2] = \|u - P_H u\|_{L^2(D)}^2, \quad (5.25)$$

since  $\dot{W}_R = \dot{W} - \dot{W}_{\mathcal{L}} = \dot{W} - P_H \dot{W}$ . Hence  $\mathcal{C}_R(u, u)$  is always non-negative and it is zero if and only if  $u \in V_H$ .  $\square$

From the proofs of lemmas 5.2 and 5.3, we see that we can interpret  $\dot{W}_{\mathcal{L}}$  as the  $L^2$ -projection of white noise onto  $V_H$ , in formulas

$$(\dot{W}_{\mathcal{L}}, v) = (P_H \dot{W}, v) := \langle \dot{W}, P_H v \rangle, \quad \forall v \in L^2(D). \quad (5.26)$$

As a consequence, the same arguments as in the proofs of lemmas 3.4 and 3.8 essentially apply and the following results hold.

**Lemma 5.4.** Let  $s > 0$ , let  $v \in H^s(D)$  and  $\mu = \min(s, 1)$ . Then

$$\mathbb{E}[|\langle \dot{W} - \dot{W}_{\mathcal{L}}, v \rangle|^2]^{1/2} \leq c(s, d, D) |\square_H|^{\mu/d} \|v\|_{H^s(D)}. \quad (5.27)$$

*Proof.* We have that, for all  $v \in L^2(D)$ ,

$$\mathbb{E}[|\langle \dot{W} - \dot{W}_{\mathcal{L}}, v \rangle|^2] = \mathbb{E}[|\langle \dot{W}, v \rangle - \langle \dot{W}, P_H v \rangle|^2] \quad (5.28)$$

$$= \mathbb{E}[|\langle \dot{W}, v - P_H v \rangle|^2] = \|v - P_H v\|_{L^2(D)}^2, \quad (5.29)$$

by the definitions of  $\dot{W}$  and the interpretation of  $\dot{W}_{\mathcal{L}}$  as  $L^2$ -projection onto  $V_H$ . As  $P_H v$  is the best approximation to  $v$  in the  $L^2(D)$  norm in  $V_H$  the result follows by standard FEM theory [21, 22].  $\square$

**Lemma 5.5.** Let  $u$  be the solution to (5.10) and let  $u_{\mathcal{L}}$  be the solution to (5.10) in which  $\dot{W}$  has been replaced with  $\dot{W}_{\mathcal{L}}$ . For any sufficiently small  $\epsilon > 0$ , if we let  $s = 2 - d/2 - \epsilon$  and  $\mu = \min(s, 1)$ , we have that for all finite  $1 \leq q \in \mathbb{N}$

$$\mathbb{E}[|u - u_{\mathcal{L}}|_{L^q(D)}^q]^{1/q} \leq \bar{c}(\epsilon, s, d, D) |\square_H|^{\mu/d}, \quad (5.30)$$

where  $\bar{c}$  blows up as  $\epsilon \rightarrow 0$  when  $\mu = s$ .

*Proof.* The proof is essentially the same as that of Lemma 3.8. Let  $g(\mathbf{x}, \mathbf{y})$  be the Green's function of equation (5.10) over  $D$ . Then we can write down analytic expressions for  $u$  and  $u_{\mathcal{L}}$  (see also [32]),

$$u = \langle \dot{W}, g(\mathbf{x}, \mathbf{y}) \rangle_x, \quad u_{\mathcal{L}} = (\dot{W}_{\mathcal{L}}, g(\mathbf{x}, \mathbf{y}))_x = \langle \dot{W}, P_H^x g(\mathbf{x}, \mathbf{y}) \rangle_x, \quad (5.31)$$

where by  $\langle \cdot, \cdot \rangle_x$ ,  $(\cdot, \cdot)_x$  and by  $P_H^x$  we mean that the pairing, inner product and orthogonal projection are taken with respect to the variable  $\mathbf{x}$ . This gives us, by linearity,

$$u - u_{\mathcal{L}} = \langle \dot{W}, g(\mathbf{x}, \mathbf{y}) - P_H^x g(\mathbf{x}, \mathbf{y}) \rangle_x = \langle \dot{W} - \dot{W}_{\mathcal{L}}, g(\mathbf{x}, \mathbf{y}) \rangle_x. \quad (5.32)$$

Fubini-Tonelli's theorem allows us to exchange the order of expectation and integration, giving us

$$\mathbb{E}[|u - u_{\mathcal{L}}|_{L^q(D)}^q] = \|\mathbb{E}[|u - u_{\mathcal{L}}|^q]\|_{L^1(D)} = \hat{c}(q) \|\mathbb{E}[|u - u_{\mathcal{L}}|^2]^{q/2}\|_{L^1(D)}, \quad (5.33)$$

where we have used the fact that the  $q$ -th moment of a zero-mean Gaussian random variable is proportional to its standard deviation to the  $q$ -th power. Lemma 5.4 then gives us

$$\mathbb{E}[|u - u_{\mathcal{L}}|_{L^q(D)}^q] = \hat{c}(q) \|\mathbb{E}[|\langle \dot{W} - \dot{W}_{\mathcal{L}}, g(\mathbf{x}, \mathbf{y}) \rangle_x|^2]^{q/2}\|_{L^1(D), y} \quad (5.34)$$

$$\leq \hat{c}(q) (c(s, d, D) |\square_H|^{\mu/d})^q \int_D \|g(\mathbf{x}, \mathbf{y})\|_{H^{\mu, x}}^q d\mathbf{y}. \quad (5.35)$$

As in the proof of Lemma 3.8, we can use elliptic regularity estimates to bound

$$\int_D \|g(\mathbf{x}, \mathbf{y})\|_{H^{\mu}(D), x}^q d\mathbf{y} \leq |D| \tilde{c}(\epsilon, s, d, D)^q, \quad (5.36)$$

where  $\tilde{c}$  blow up as  $\epsilon \rightarrow 0$  when  $\mu = s$  (in 2D and 3D). In 1D the Green's function is more regular than  $H^1(D)$  so the constant does not blow up. We finally conclude

$$\mathbb{E}[||u - u_{\mathcal{L}}||_{L^q(D)}^q]^{1/q} \leq \bar{c}(\epsilon, q, s, d, D) |\square_H|^{\mu/d}, \quad (5.37)$$

with  $\bar{c} = |D|^{1/q} c \hat{c}^{1/q} \tilde{c}$ , which is the proposition.  $\square$

*Remark 5.10.* Thanks to these results we deduce that if  $D_{\mathcal{L}}$  is fine enough (or if  $V_h \equiv V_H$ ), the correction  $\dot{W}_R$  is not needed at all. However, Haar wavelets are only piecewise constant and we can only expect at most first order convergence of  $\dot{W}_{\mathcal{L}}$  to  $\dot{W}$ : large QMC dimensions and an extremely fine Haar mesh are needed to make the correction term  $\dot{W}_R$  negligible and this translates into very expensive samples of  $\dot{W}_{\mathcal{L}}$ . If we also compute samples of  $\dot{W}_R$ , however, the Haar level can be kept small.

The sampling of  $\dot{W}_R$  can be performed independently on each Haar cell. If we focus our attention only on the basis functions  $\phi_1, \dots, \phi_{m_k} \in V_h$  of support entirely contained within a given Haar cell  $\square_k$ , we note that the expression (5.23) for  $\mathcal{C}_R$  simplifies to

$$\mathcal{C}_R(\phi_i, \phi_j) = (\phi_i, \phi_j) - \frac{1}{|\square_k|} \int_{\square_k} \phi_i \, d\mathbf{x} \int_{\square_k} \phi_j \, d\mathbf{x}, \quad \text{for all } i, j \in \{1, \dots, m_k\}. \quad (5.38)$$

Similarly as in Chapter 4, the sampling of  $\dot{W}_R$  over  $\square_k$  boils down to sampling a zero-mean Gaussian vector  $\mathbf{b}_R^k$  with entries  $(\mathbf{b}_R^k)_i = \langle \dot{W}_R, \phi_i \rangle$  and with covariance matrix  $C_R^k$  of entries  $(C_R^k)_{ij}$  given by

$$\mathbf{b}_R^k \sim \mathcal{N}(0, C_R^k), \quad (C_R^k)_{ij} = \mathcal{C}_R(\phi_i, \phi_j). \quad (5.39)$$

If we let  $M_k$  be the local mass matrix over the space spanned by the  $\{\phi_i\}_{i=1}^{m_k}$ , with entries  $(M_k)_{ij} = (\phi_i, \phi_j)$  and if we let the vector  $\mathbf{I}^k \in \mathbb{R}^{m_k}$  be given by

$$\mathbf{I}^k = \left[ \int_{\square_k} \phi_1 \, d\mathbf{x}, \dots, \int_{\square_k} \phi_{m_k} \, d\mathbf{x} \right]^T, \quad (5.40)$$

we can write  $C_R^k$  as

$$C_R^k = M_k - \frac{1}{|\square_k|} \mathbf{I}^k (\mathbf{I}^k)^T. \quad (5.41)$$

Note that a consequence of Lemma 5.3 is that  $C_R^k$  is positive semi-definite with null-space spanned by the vector  $\mathbf{1} \in \mathbb{R}^{m_k}$ , the length  $m_k$  vector of all ones (piecewise constant functions over  $D_{\mathcal{L}}$  are constant over each Haar cell and are in the null-space of the covariance). The sampling of a Gaussian vector with this covariance through factorisation is expensive as direct factorisation of  $C_R^k$  (e.g. truncated Cholesky [84] or singular value decomposition) has an  $O(m_k^3)$  and  $O(m_k^2)$  cost and memory complexity respectively and it is therefore to be avoided.

We now show how  $\mathbf{b}_R^k$  can be sampled efficiently by extending the techniques presented in Chapter 4. The main idea is to first sample a Gaussian vector with

covariance  $M_k$  in linear complexity and then perform an efficient update to obtain a sample of  $\dot{\mathbf{W}}_R$ . We can write the action of  $\dot{\mathbf{W}}_R$  against each  $\phi_i$  as

$$\langle \dot{\mathbf{W}}_R, \phi_i \rangle = \langle \dot{\mathbf{W}} - \dot{\mathbf{W}}_{\mathcal{L}}, \phi_i \rangle = \langle \dot{\mathbf{W}}, \phi_i \rangle - \langle \dot{\mathbf{W}}_{\mathcal{L}}, \phi_i \rangle = (\mathbf{b}_M^k)_i - w_k(\mathbf{I}^k)_i, \quad (5.42)$$

where  $\mathbf{I}^k$  is given by (5.40),  $w_k$  by

$$w_k = \frac{1}{|\square_k|} \langle \dot{\mathbf{W}}, \mathbf{1}_{\square_k} \rangle, \quad w_k \sim \mathcal{N}\left(0, \frac{1}{|\square_k|}\right), \quad (5.43)$$

and the vector  $\mathbf{b}_M^k \in \mathbb{R}^{m_k}$  is given entrywise by

$$(\mathbf{b}_M^k)_i = \langle \dot{\mathbf{W}}, \phi_i \rangle, \quad i = 1, \dots, m_k. \quad (5.44)$$

The variables  $w_k$  and  $\mathbf{b}_M^k$  are by Definition 2.6 all zero-mean joint Gaussian variables with covariance

$$\mathbb{E}[w_k w_k] = \frac{1}{|\square_k|}, \quad \mathbb{E}[\mathbf{b}_M^k w_k] = \frac{\mathbf{I}^k}{|\square_k|}, \quad \mathbb{E}[\mathbf{b}_M^k (\mathbf{b}_M^k)^T] = M_k. \quad (5.45)$$

Thanks to these relations and to (5.42), if we set

$$\mathbf{b}_R^k = \mathbf{b}_M^k - w_k \mathbf{I}^k, \quad (5.46)$$

then the covariance of  $\mathbf{b}_R^k$  is correct (cf. equation (5.41)) since

$$\begin{aligned} \mathbb{E}[\mathbf{b}_R^k (\mathbf{b}_R^k)^T] &= \mathbb{E}[(\mathbf{b}_M^k - w_k \mathbf{I}^k)(\mathbf{b}_M^k - w_k \mathbf{I}^k)^T] \\ &= \mathbb{E}[\mathbf{b}_M^k (\mathbf{b}_M^k)^T] - \mathbb{E}[\mathbf{b}_M^k w_k](\mathbf{I}^k)^T - \mathbf{I}^k \mathbb{E}[\mathbf{b}_M^k w_k]^T + \mathbb{E}[w_k w_k] \mathbf{I}^k (\mathbf{I}^k)^T \\ &= M_k - \frac{1}{|\square_k|} \mathbf{I}^k (\mathbf{I}^k)^T. \end{aligned} \quad (5.47)$$

In what follows, we assume that constants can be represented exactly by the FEM subspace  $V_h$ , i.e.  $c \in V_h$  for all  $c \in \mathbb{R}$ . This assumption is satisfied by most commonly used finite element subspaces (e.g. Lagrange). Let  $\phi_k = [\phi_1, \dots, \phi_{m_k}]^T$ . This means that for each Haar cell  $\square_k$  there exists a vector  $\mathbf{c}_k \in \mathbb{R}^{m_k}$  such that  $\mathbf{1}_{\square_k} \equiv \mathbf{c}_k \cdot \phi_k$ . It is then straightforward to obtain  $w_k$  from  $\mathbf{b}_M^k$  since

$$\mathbf{c}_k \cdot \mathbf{b}_M^k = \sum_{i=1}^{m_k} \langle \dot{\mathbf{W}}, (\mathbf{c}_k)_i \phi_i \rangle = \langle \dot{\mathbf{W}}, \mathbf{c}_k \cdot \phi_k \rangle = \langle \dot{\mathbf{W}}, \mathbf{1}_{\square_k} \rangle = |\square_k| w_k, \quad (5.48)$$

hence  $w_k = |\square_k|^{-1} \mathbf{c}_k \cdot \mathbf{b}_M^k$ . Note that  $\mathbf{c}_k$  is always known, e.g. for Lagrange basis functions on simplices we have  $\mathbf{c}_k = \mathbf{1} \in \mathbb{R}^{m_k}$ .

We can now sample  $\dot{\mathbf{W}}_R$  from its distribution by using the following algorithm, in which we exploit the same strategy presented in Chapter 4:

### Algorithm for the efficient sampling of $\dot{W}_R$ .

1. Loop over each Haar cell  $\square_k$ .
2. Use the technique presented in Chapter 4 for the independent case to work supermesh cell by supermesh cell and sample a Gaussian vector  $\mathbf{b}_M^k \sim \mathcal{N}(0, M_k)$  in linear cost complexity.
3. Set  $w_k = |\square_k|^{-1} \mathbf{c}_k \cdot \mathbf{b}_M^k$  and compute  $\mathbf{b}_R^k = \mathbf{b}_M^k - w_k \mathbf{I}^k$ .

*Remark 5.11.* This algorithm has  $O(\mathcal{N}_S m_e^3)$  cost and  $O(\mathcal{N}_S m_e^2)$  memory complexity, where  $\mathcal{N}_S$  is the total number of supermesh cells (cf. Chapter 4). As in Remark 5.8 we expect  $\mathcal{N}_S$  to be in practice of  $O(\mathcal{N}_h + \mathcal{N}_{\mathcal{L}})$ , where  $\mathcal{N}_h$  and  $\mathcal{N}_{\mathcal{L}}$  are the number of cells of  $D_h$  and of  $D_{\mathcal{L}}$  respectively.

*Remark 5.12* (QMC in the affine case). Let  $V_S$  be a FEM subspace of the same family of  $V_h$ , but defined over the supermesh  $S_h$ . In what we have presented so far we showed how to sample the action of  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$  against functions in  $V_h$ . An alternative strategy that is particularly advantageous in the case in which the mapping to the reference cell of the FEM basis chosen is affine (e.g. Lagrange elements on simplices) consists of computing the action of  $\dot{W}_{\mathcal{L}}$  and  $\dot{W}_R$  against basis functions of the supermesh space  $V_S$  instead and then transfer (by interpolation or projection) the result onto  $V_h$ . This transfer operation can be performed in linear time and it is exact since  $V_h$  is nested within  $V_S$  by construction. The advantage in doing this is that, in the affine case, as shown in Chapter 4, for each supermesh cell  $e$ , the local mass matrices  $M_e$  and local vectors  $\mathbf{I}_e = [\int_D \phi_1 \, d\mathbf{x}, \dots, \int_D \phi_{m_e} \, d\mathbf{x}]^T$  of the space  $V_S$  are constant up to a scalar volume factor. This allows us to save a great deal of computations by pre-computing a single  $\mathbf{I}_e$  vector and a single Cholesky factorisation of a  $M_e$  (see remark 4.2 and corollary 4.3 in [43]).

## 5.4 A partial QMC convergence result

The white noise sampling strategy we just presented is hybrid in the sense that both a randomised QMC point rule and a pseudo-random sequences are used. It is then unclear what the order of convergence with respect to the number of samples should be. The hope is, of course, to achieve something better than the standard MC rate of convergence or otherwise there would be no use for the QMC method presented. Proving convergence results with respect to the number of samples for QMC (and MLQMC) in a PDE setting is non-trivial and results have so far been established only for a limited class of QMC point sequences [88, 109].

Although deriving a convergence estimate is outside of the scope of this work, in what follows we try to build up intuition about what is likely to be happening in practice. As an example problem, consider the following elliptic PDE with log-normal diffusion coefficient (already in weak form): find  $p(\cdot, \omega) \in H_0^1(G)$  such that for all  $v \in H_0^1(G)$ ,

$$a(p, v) = (D^*(\mathbf{x}, \omega) \nabla p, \nabla v) = (f, v), \quad \text{a.s.}, \quad D^*(\mathbf{x}, \omega) = e^{u(\mathbf{x}, \omega)}, \quad (5.49)$$

where for a given sampling domain  $D \subset\subset \mathbb{R}^d$  we have that  $G \subset\subset D$  is a domain of class  $C^{1,\epsilon}$  for any  $\epsilon > 0$ ,  $f \in L^\infty(G)$  (these hypotheses imply  $p \in C^{1,\epsilon}(\bar{G})$  a.s. for any  $\epsilon > 0$ , see theorem 8.34 in [66]) and  $u(\mathbf{x}, \omega)$  is a Matérn field satisfying equation (5.10) over  $D$ , and assume that we are interested in computing the expectation of a possibly nonlinear output functional of  $p$ , namely  $P(p)$ . We will now establish the following result:

**Theorem 5.6.** *Let  $p \in W^{1,\infty}(G)$  a.s. be the solution of (5.10) where  $u \in H^s(G)$  a.s. has been sampled using the hybrid QMC technique presented in this chapter. Let  $u_{\mathcal{L}}$  be the solution of (5.10) obtained by using the same  $\dot{W}_{\mathcal{L}}$  sample as for  $u$  and by setting  $\dot{W}_R = 0$ , and let  $p_{\mathcal{L}}(\cdot, \omega) \in H_0^1(G)$  satisfy, for all  $v \in H_0^1(G)$ ,*

$$a_{\mathcal{L}}(p_{\mathcal{L}}, v) = (D_{\mathcal{L}}^*(\mathbf{x}, \omega) \nabla p_{\mathcal{L}}, \nabla v) = (f, v), \quad \text{a.s.}, \quad D_{\mathcal{L}}^*(\mathbf{x}, \omega) = e^{u_{\mathcal{L}}(\mathbf{x}, \omega)}. \quad (5.50)$$

Let  $\mu = \min(s, 1)$  and let  $\mathcal{L}$  be the Haar level used to sample  $u$ . Assume that the functional  $P$  is continuously Fréchet differentiable and let  $\hat{P}$ ,  $\hat{P}_{\mathcal{L}}$  and  $\widehat{P - P_{\mathcal{L}}}$  be the QMC estimators for  $\mathbb{E}[P]$ ,  $\mathbb{E}[P_{\mathcal{L}}]$  and  $\mathbb{E}[P - P_{\mathcal{L}}]$  respectively, obtained by using  $N$  QMC points. Here,  $P = P(p)$  and  $P_{\mathcal{L}} = P(p_{\mathcal{L}})$ . If there exist constants  $c > 0$  and  $q \geq 1$ ,  $N_0 > 1$  such that the QMC estimators satisfy for  $N > N_0$ ,

$$\mathbb{V}[\hat{P}] \leq c \frac{\mathbb{V}[P]}{N}, \quad \mathbb{V}[\hat{P}_{\mathcal{L}}] \leq c \frac{\mathbb{V}[P_{\mathcal{L}}]}{N^q}, \quad \mathbb{V}[\widehat{P - P_{\mathcal{L}}}] \leq c \frac{\mathbb{V}[P - P_{\mathcal{L}}]}{N}, \quad (5.51)$$

i.e. the QMC estimators are never asymptotically worse than standard Monte Carlo, then there exists a sufficiently large constants  $\mathcal{L}_0$  such that the statistical error  $\mathbb{V}[\hat{P}]$  also satisfies for  $N > N_0$ ,  $\mathcal{L} > \mathcal{L}_0$ ,

$$\mathbb{V}[\hat{P}] \leq \frac{c}{N^q} + \frac{\bar{c}}{N} 2^{-\mu \mathcal{L}}, \quad (5.52)$$

where  $\bar{c} > 0$  is independent from  $\mathcal{L}$  and  $N$ .

*Remark 5.13.* Condition (5.51) is satisfied by most randomised QMC point sets, e.g. randomised digital nets and sequences and randomly shifted lattice rules [112, 134, 135, 136].

*Proof.* In this proof, we work pathwise for fixed  $\omega \in \Omega$ . We start with essentially the same duality argument that yields lemma 3.2 in [166]. Let  $v, \bar{v} \in H^1(G)$  and let  $D_v P(\bar{v})$  be the Gateaux derivative of  $P$  at  $\bar{v}$ , namely

$$D_v P(\bar{v}) := \lim_{\epsilon \rightarrow 0} \frac{P(\bar{v} + \epsilon v) - P(\bar{v})}{\epsilon}, \quad (5.53)$$

define the average derivative of  $P$  on the path from  $p$  to  $p_{\mathcal{L}}$ ,

$$\overline{D_v P}(p, p_{\mathcal{L}}) = \int_0^1 D_v P(p + \theta(p_{\mathcal{L}} - p)) \, d\theta, \quad (5.54)$$

and introduce the dual problem: find  $z(\cdot, \omega) \in H_0^1(G)$  s.t.

$$a(v, z) = \overline{D_v P}(p, p_{\mathcal{L}}), \quad \forall v \in H_0^1(G). \quad (5.55)$$

The fundamental theorem of calculus for Fréchet derivatives then yields,

$$P - P_{\mathcal{L}} = \int_0^1 D_{p-p_{\mathcal{L}}} P(p + \theta(p_{\mathcal{L}} - p)) \, d\theta = \overline{D_{p-p_{\mathcal{L}}} P}(p, p_{\mathcal{L}}) = a(p - p_{\mathcal{L}}, z), \quad \text{a.s.} \quad (5.56)$$

Applying Hölder's inequality gives,

$$|P - P_{\mathcal{L}}| = |a(p - p_{\mathcal{L}}, z)| \leq \|D^*\|_{L^\infty(G)}(\omega) |z|_{H^1(G)} |p - p_{\mathcal{L}}|_{H^1(G)} \quad \text{a.s.} \quad (5.57)$$

We now need a bound for  $|p - p_{\mathcal{L}}|_{H^1(G)}$ . Note that, a.s. for all  $v \in H_0^1(G)$ ,

$$0 = a(p, v) - a_{\mathcal{L}}(p_{\mathcal{L}}, v) = a(p - p_{\mathcal{L}}, v) + a(p_{\mathcal{L}}, v) - a_{\mathcal{L}}(p_{\mathcal{L}}, v). \quad (5.58)$$

Setting  $v = p - p_{\mathcal{L}}$  gives

$$0 \leq a(p - p_{\mathcal{L}}, p - p_{\mathcal{L}}) = a_{\mathcal{L}}(p_{\mathcal{L}}, v) - a(p_{\mathcal{L}}, v) = ((D_{\mathcal{L}}^* - D^*) \nabla p_{\mathcal{L}}, \nabla(p - p_{\mathcal{L}})) \quad \text{a.s.} \quad (5.59)$$

Both quantities can be bounded as follows: let  $D_{\min}^*(\omega) = \inf_{x \in D} |D^*(\cdot, \omega)|$ ,

$$0 \leq D_{\min}^* |p - p_{\mathcal{L}}|_{H_0^1(G)}^2 \leq a(p - p_{\mathcal{L}}, p - p_{\mathcal{L}}) \quad (5.60)$$

$$((D_{\mathcal{L}}^* - D^*) \nabla p_{\mathcal{L}}, \nabla(p - p_{\mathcal{L}})) \leq \|D^* - D_{\mathcal{L}}^*\|_{L^2(G)} \|\nabla p_{\mathcal{L}}\|_{L^\infty(G)} |p - p_{\mathcal{L}}|_{H^1(G)}, \quad (5.61)$$

almost surely, hence yielding, after division by  $|p - p_{\mathcal{L}}|_{H^1(G)}$ ,

$$|p - p_{\mathcal{L}}|_{H^1(G)} \leq \beta \frac{\|D^*\|_{L^\infty(G)}^\alpha}{D_{\min}^*} \|\nabla p\|_{L^\infty(G)} \|u - u_{\mathcal{L}}\|_{L^2(G)}, \quad \text{a.s.,} \quad (5.62)$$

for some  $\alpha, \beta \geq 1$ . Here we used the fact that since  $u_{\mathcal{L}}$  converges to  $u$  in  $C^{2-d/2-\epsilon}(D)$  for any  $\epsilon > 0$  [42] (and consequently in  $C^{2-d/2-\epsilon}(G)$ ), there exists a Haar level  $\mathcal{L}_0$  such that for all  $\mathcal{L} > \mathcal{L}_0$  there exists a constant  $\alpha \geq 1$  independent from  $\mathcal{L}$  such that  $\|u_{\mathcal{L}}\|_{L^\infty(G)} \leq \alpha \|u\|_{L^\infty(G)}$ , hence

$$\|D^* - D_{\mathcal{L}}^*\|_{L^2(G)} \leq e^{\max(\|u\|_{L^\infty(G)}, \|u_{\mathcal{L}}\|_{L^\infty(G)})} \|u - u_{\mathcal{L}}\|_{L^2(G)} \quad (5.63)$$

$$\leq \|D^*\|_{L^\infty(G)}^\alpha \|u - u_{\mathcal{L}}\|_{L^2(G)}, \quad \text{a.s.} \quad (5.64)$$

Similarly, we also used that for sufficiently large  $\mathcal{L}$ , we have  $\|\nabla p_{\mathcal{L}}\|_{L^\infty(G)} \leq \beta \|\nabla p\|_{L^\infty(G)}$  for some  $\beta \geq 1$  since  $p_{\mathcal{L}}$  converges to  $p$  in  $W^{1,\infty}(G)$ . Putting (5.57) and (5.62) together yields,

$$|P - P_{\mathcal{L}}| \leq C(\omega) \|u - u_{\mathcal{L}}\|_{L^2(G)}, \quad \text{a.s.,} \quad (5.65)$$

where  $C(\omega)$  and is given by

$$C(\omega) = \beta \frac{\|D^*\|_{L^\infty(G)}^{\alpha+1}}{D_{\min}^*} \|\nabla p\|_{L^\infty(G)} |z|_{H^1(G)}. \quad (5.66)$$

Note that since all terms involved are in  $L^t(\Omega, \mathbb{R})$  for all  $t \in (0, \infty)$  (see proposition 2.4 and theorem 3.4 in [39]), we have that  $C(\omega) \in L^t(\Omega, \mathbb{R})$  for all  $t \in [1, \infty)$  due to the generalised Hölder inequality. We now note that

$$\begin{aligned} \mathbb{V}[\hat{P}] &= \mathbb{V}[\hat{P}_\mathcal{L} + \hat{P} - \hat{P}_\mathcal{L}] = \mathbb{V}[\hat{P}_\mathcal{L}] + \text{Cov}(\hat{P} + \hat{P}_\mathcal{L}, \hat{P} - \hat{P}_\mathcal{L}) \\ &\leq \mathbb{V}[\hat{P}_\mathcal{L}] + \mathbb{V}[\hat{P} + \hat{P}_\mathcal{L}]^{1/2} \mathbb{V}[(\hat{P} - \hat{P}_\mathcal{L})]^{1/2} \end{aligned} \quad (5.67)$$

$$\leq c \frac{\mathbb{V}[P_\mathcal{L}]}{N^q} + \frac{c\tilde{c}}{N} \mathbb{V}[P]^{1/2} \mathbb{V}[P - P_\mathcal{L}]^{1/2}. \quad (5.68)$$

where we used the Cauchy-Schwarz inequality, hypothesis (5.51) and the fact that since  $P$  converges to  $P_\mathcal{L}$  as  $\mathcal{L} \rightarrow \infty$  a.s. and in  $L^t(\Omega, \mathbb{R})$  for  $t \in (1, \infty)$ , there exists a Haar level  $\mathcal{L}_0$  s.t. for all  $\mathcal{L} > \mathcal{L}_0$  there exists  $c_1 > 0$  s.t.  $\mathbb{V}[P_\mathcal{L}] \leq c_1 \mathbb{V}[P]$  and consequently another constant  $\tilde{c} > 0$  s.t.  $\mathbb{V}[P + P_\mathcal{L}]^{1/2} \leq \tilde{c} \mathbb{V}[P]^{1/2}$ . Combining equation (5.65) with Cauchy-Schwarz and the embedding  $L^4(G) \subset L^2(G)$  gives that

$$\mathbb{V}[P - P_\mathcal{L}]^{1/2} \leq |D|^{1/4} \mathbb{E}[C(\omega)^4]^{1/4} \mathbb{E}[||u - u_\mathcal{L}||_{L^4(G)}^4]^{1/4}. \quad (5.69)$$

Now, owing to Lemma 5.5, we have that

$$\mathbb{E}[||u - u_\mathcal{L}||_{L^4(G)}^4]^{1/4} \leq \bar{C}(s, D, d) |\square_H|^{\mu/d}, \quad (5.70)$$

for some constant  $\bar{C}$ . Note that by construction  $|\square_H| = 2^{-d\mathcal{L}}$ . We can now pull together equations (5.67), (5.69) and (5.70) to obtain,

$$\mathbb{V}[\hat{P}] \leq c \frac{\mathbb{V}[P_\mathcal{L}]}{N^q} + \frac{\bar{c}}{N} 2^{-\mu\mathcal{L}}, \quad (5.71)$$

where

$$\bar{c} = c\tilde{c}|D|^{1/4} \bar{C}(s, D, d) \mathbb{E}[C(\omega)^4]^{1/4} \mathbb{V}[P]^{1/2}. \quad (5.72)$$

and this concludes the proof.  $\square$

Theorem 5.6 states that the statistical error introduced by approximating  $\mathbb{E}[P]$  with our hybrid QMC technique can be split in two terms, where the former is the statistical error of a pure randomised QMC estimator and might converge faster than  $O(N^{-1/2})$  and the second is a standard MC error correction term that exhibits the usual Monte Carlo rate, but decays geometrically as the Haar level increases. This splitting of the error directly relates to the splitting of white noise as the first term in 5.52 only depends on the truncation  $\dot{W}_\mathcal{L}$ . If  $\dot{W}_\mathcal{L}$  well approximates  $\dot{W}$ , then we expect a pure QMC rate in the pre-asymptotic regime, while if the approximation is poor (small  $\mathcal{L}$ ), then only a  $O(N^{-1/2})$  rate can be expected.

Another way of interpreting our hybrid approach is that we are splitting white noise into a smooth part  $\dot{W}_{\mathcal{L}}$  and a rough part  $\dot{W}_R$ . QMC is effective at reducing the statistical error coming from the smooth part, but performs poorly when approximating the rough part and we are better off with directly using pseudo-random points. This aspect can be seen as another instance of the *effective dimensionality* principle that we introduced in Section 2.1 and it was experimentally investigated by Beentjes and Baker in [19]. Beentjes and Baker observe that when applying QMC to integrands that can be separated into a smooth and a rough component an initial QMC-like rate is seen, which asymptotically decays into a standard Monte Carlo rate as the number of samples is increased. Beentjes and Baker also observe that integrands with smaller rough components will transition later to the lower  $O(N^{-1/2})$  convergence rate. Similarly, Theorem 5.6 states that the larger  $\mathcal{L}$  is, the smaller the rough correction component  $\dot{W}_R$  will be and the later the fast  $O(N^{-q})$  rate will switch to a slow  $O(N^{-1/2})$  rate.

As we will see further on in this chapter, even if the asymptotic rate is still  $O(N^{-1/2})$ , large gains are still obtained in practice in the pre-asymptotic QMC-like regime, especially in a MLQMC setting (presented next) where not that many samples are needed on the finest levels.

## 5.5 Sampling coupled realisations for MLQMC

We now generalise the QMC sampling algorithm just presented to the MLQMC case. Compared to standard Monte Carlo, both MLMC and QMC already bring a significant computational improvement. When the two are combined into MLQMC, it is sometimes possible to obtain the best of two worlds and further improve the computational complexity and speed. However, to do so, we must be able to satisfy the requirements and assumptions underlying both QMC and MLMC: we must order the dimensions of our random input in decaying order of importance as in QMC and introduce an approximation level hierarchy and enforce a good coupling between the levels as in MLMC. We now show how this can be done with white noise sampling.

In what follows we assume we have a MLQMC hierarchy of possibly non-nested FEM approximation subspaces  $\{V^\ell\}_{\ell=1}^L$  over the meshes  $\{D_h^\ell\}_{\ell=1}^L$  and of accuracy increasing with  $\ell$ . Since as in the MLMC case (see Chapter 4) the only stochastic element in (5.10) is white noise, on each MLQMC level we must be able to draw Matérn field samples  $u_\ell \in V^\ell$  and  $u_{\ell-1} \in V^{\ell-1}$  for  $\ell > 1$  that satisfy the following variational problems coupled by the same white noise sample: for a given  $\omega_\ell^n \in \Omega$ , find  $u_\ell \in V^\ell$  and  $u_{\ell-1} \in V^{\ell-1}$  such that

$$(u_\ell, v_\ell) + \kappa^{-2}(\nabla u_\ell, \nabla v_\ell) = \langle \dot{W}, v_\ell \rangle(\omega_\ell^n), \quad \text{for all } v_\ell \in V^\ell, \quad (5.73)$$

$$(u_{\ell-1}, v_{\ell-1}) + \kappa^{-2}(\nabla u_{\ell-1}, \nabla v_{\ell-1}) = \langle \dot{W}, v_{\ell-1} \rangle(\omega_\ell^n), \quad \text{for all } v_{\ell-1} \in V^{\ell-1}. \quad (5.74)$$

where the terms on the right-hand side are coupled in the sense that they are centred Gaussian random variables with covariance  $\mathbb{E}[\langle \dot{W}, v_s \rangle \langle \dot{W}, v_s \rangle] = (v_s, v_s)$  for  $s, s \in \{\ell, \ell-1\}$ , as given by Definition 2.6. Again we order the dimensions of white noise

by expanding it in the Haar wavelet basis as in (5.6), but this time we allow the Haar level to possibly increase with the MLQMC level and we split the expansion at the finer Haar level between the two MLQMC levels,  $\mathcal{L}_\ell$ ,

$$\dot{\mathbf{W}} = \dot{\mathbf{W}}_{\mathcal{L}_\ell} + \dot{\mathbf{W}}_{R_\ell}, \quad (5.75)$$

where the splitting of the expansion is done in the same way as in equation (5.8). The choice of Haar levels  $\mathcal{L}_\ell$  on each MLQMC level is problem-dependent. We discuss this further in Section 5.6. From now on we assume that  $\mathcal{L}_{\ell-1} \leq \mathcal{L}_\ell$ , although extending the methods presented to decreasing Haar level hierarchies as well is straight-forward. Note that the splitting of the expansion at the MLQMC level  $\ell$  is done at Haar level  $\mathcal{L}_\ell$  independently from the value of  $\mathcal{L}_{\ell-1}$ , i.e.  $\dot{\mathbf{W}}$  on the coarser grid is always sampled with Haar level  $\mathcal{L}_\ell$  on MLQMC level  $\ell$  and not with Haar level  $\mathcal{L}_{\ell-1}$ . However, this does not affect the MLQMC telescoping sum since  $\dot{\mathbf{W}}$  is always sampled without bias on all levels owing to the correction term  $\dot{\mathbf{W}}_{R_\ell}$ .

Let  $\{\phi_i^\ell\}_{i=1}^{m_\ell}$  and  $\{\phi_j^{\ell-1}\}_{j=1}^{m_{\ell-1}}$  be the basis functions spanning  $V^\ell$  and  $V^{\ell-1}$  respectively. Sampling white noise on both MLQMC levels again means to sample the vectors  $\mathbf{b}_{\mathcal{L}}^\ell$ ,  $\mathbf{b}_{\mathcal{L}}^{\ell-1}$ ,  $\mathbf{b}_R^\ell$  and  $\mathbf{b}_R^{\ell-1}$ , with entries given by,

$$(\mathbf{b}_{\mathcal{L}}^s)_i = \langle \dot{\mathbf{W}}_{\mathcal{L}_\ell}, \phi_i^s \rangle, \quad (\mathbf{b}_R^s)_i = \langle \dot{\mathbf{W}}_{R_\ell}, \phi_i^s \rangle, \quad \text{for } i = 1, \dots, m_s, \quad s \in \{\ell, \ell-1\}. \quad (5.76)$$

Since we both require a multilevel coupling and a Haar wavelet expansion, this time we need to construct a *three-way* supermesh  $S_h$  between  $D_{\mathcal{L}_\ell}$ ,  $D_h^\ell$  and  $D_h^{\ell-1}$  (note that  $D_{\mathcal{L}_{\ell-1}}$  is always nested within  $D_{\mathcal{L}_\ell}$  so there is no need for a four-way supermesh). Thanks to the supermesh construction we can split the support of all the basis functions so that each  $\phi_i^\ell$  and  $\phi_j^{\ell-1}$  has support entirely contained within a single Haar cell. In fact, we will assume for simplicity from now on that the supports of all basis functions have this property. The sampling of  $\dot{\mathbf{W}}$  in the MLQMC case is extremely similar to that of the QMC case with only a few differences concerning the sampling of  $\dot{\mathbf{W}}_{R_\ell}$  which we will now highlight.

Again, portions of  $\dot{\mathbf{W}}_{R_\ell}$  on separate Haar cells of  $D_{\mathcal{L}_\ell}$  are independent and we can therefore sample  $\dot{\mathbf{W}}_{R_\ell}$  Haar cell-wise. For each Haar cell  $\square_k$  and for  $s \in \{\ell, \ell-1\}$ , let  $\phi_1^s, \dots, \phi_{m_k^s}^s$  be the basis functions with non-zero support over  $\square_k$  and define the Haar cell correction vectors  $\mathbf{b}_{R,k}^s$  with entries  $(\mathbf{b}_{R,k}^s)_i = \langle \dot{\mathbf{W}}, \phi_i^s \rangle$  for  $i \in \{1, \dots, m_k^s\}$  and covariances given by,

$$\mathbb{E}[\mathbf{b}_{R,k}^s (\mathbf{b}_{R,k}^s)^T] = M_k^s - \frac{1}{|\square_k|} \mathbf{I}_s^k (\mathbf{I}_s^k)^T, \quad \mathbb{E}[\mathbf{b}_{R,k}^\ell (\mathbf{b}_{R,k}^{\ell-1})^T] = M_k^{\ell, \ell-1} - \frac{1}{|\square_k|} \mathbf{I}_\ell^k (\mathbf{I}_{\ell-1}^k)^T, \quad (5.77)$$

where  $(M_k^s)_{ij} = (\phi_i^s, \phi_j^s)$ ,  $(M_k^{\ell, \ell-1})_{ij} = (\phi_i^\ell, \phi_j^{\ell-1})$  and  $(\mathbf{I}_s^k)_i = \int_D \phi_i^s \, d\mathbf{x}$ . If we define  $w_k$  as in (5.43) we can again write

$$\mathbf{b}_{R,k}^s = \mathbf{b}_{M,k}^s - w_k \mathbf{I}_s^k, \quad \text{for } s \in \{\ell, \ell-1\}, \quad (5.78)$$

where  $\mathbf{b}_{M,k}^s \sim \mathcal{N}(0, M_k^s)$ . If we assume that constants can be represented exactly by both  $V^\ell$  and  $V^{\ell-1}$ , i.e. that for all  $c \in \mathbb{R}$  and for all  $s \in \{\ell, \ell-1\}$ , we have that  $c \in V^s$ ,

then there exist two vectors  $\mathbf{c}_k^\ell$  and  $\mathbf{c}_k^{\ell-1}$  such that  $\mathbf{1}_{\square_k} \equiv \mathbf{c}_k^\ell \cdot \boldsymbol{\phi}_k^\ell \equiv \mathbf{c}_k^{\ell-1} \cdot \boldsymbol{\phi}_k^{\ell-1}$ , where  $\boldsymbol{\phi}_k^s = [\phi_1^s, \dots, \phi_{m_k^s}^s]^T$  for  $s \in \{\ell, \ell-1\}$ . The same argument used to derive equation (5.48) then gives

$$w_k = \frac{1}{|\square_k|} \mathbf{c}_k^\ell \cdot \mathbf{b}_{M,k}^\ell = \frac{1}{|\square_k|} \mathbf{c}_k^{\ell-1} \cdot \mathbf{b}_{M,k}^{\ell-1}. \quad (5.79)$$

We can now proceed with the coupled sampling of  $\dot{W}$  for MLQMC as follows:

### Algorithm for the efficient sampling of $\dot{W}$ for MLQMC.

1. Compute the three-way supermesh between the FEM meshes and the Haar mesh  $D_{\mathcal{L}_\ell}$  and split the support of the basis functions of  $V^\ell$  and  $V^{\ell-1}$  to obtain  $\{\phi_i^s\}_{i=1}^{m^s}$  for  $s \in \{\ell, \ell-1\}$  each of which with support entirely contained within a single Haar cell. Compute the scalar maps  $\kappa^s(i)$  that map each  $i$  to the index  $k$  of the Haar cell  $\square_k$  that contains the support of  $\phi_i^s$  and compute  $\int_D \phi_i^s d\mathbf{x}$  for all  $i = 1, \dots, m^\ell$  and for  $s \in \{\ell, \ell-1\}$ . This step can be done once and for all as an offline step.
2. Let  $\mathcal{N}_{\mathcal{L}_\ell}$  be the number of cells of  $D_{\mathcal{L}_\ell}$ . Sample the vector  $\mathbf{z}_{\mathcal{L}_\ell} \in \mathbb{R}^{\mathcal{N}_{\mathcal{L}_\ell}}$  of the coefficients in the expression for  $\dot{W}_{\mathcal{L}_\ell}$  as  $\mathbf{z}_{\mathcal{L}} = [\mathbf{z}_{\text{QMC}}^T, \mathbf{z}_{\text{MC}}^T]^T$ , where  $\mathbf{z}_{\text{QMC}}$  is a randomised QMC point of dimension equal to the number of coefficients with  $|\mathbf{l}|_1 \leq \mathcal{L}_\ell$  and  $\mathbf{z}_{\text{MC}}$  is sampled with a pseudo-random number generator.
3. Compute the Haar cell values  $\bar{w}_k$  of  $\dot{W}_{\mathcal{L}}$  over all  $\square_k$  for  $k = 1, \dots, \mathcal{N}_{\mathcal{L}_\ell}$  in the same way as in the QMC case (this step does not depend on the FEM meshes).
4. Use the technique presented in Chapter 4 for the coupled case to work supermesh cell by supermesh cell and sample in linear cost complexity the coupled Gaussian vectors  $\mathbf{b}_{M,k}^\ell$  and  $\mathbf{b}_{M,k}^{\ell-1}$  with covariance,

$$\left[ \begin{array}{c|c} \mathbf{b}_{M,k}^\ell & \\ \hline \mathbf{b}_{M,k}^{\ell-1} & \end{array} \right] = \left[ \begin{array}{c|c} M_k^\ell & M_k^{\ell, \ell-1} \\ \hline (M_k^{\ell, \ell-1})^T & M_k^{\ell-1} \end{array} \right]. \quad (5.80)$$

5. For all  $s \in \{\ell, \ell-1\}$ , compute  $(\mathbf{b}_{\mathcal{L}})_i = \bar{w}_{\kappa^s(i)} \int_D \phi_i^s d\mathbf{x}$  for all  $i = 1, \dots, m^s$ , then set  $w_k = |\square_k|^{-1} \mathbf{c}_k^{\ell-1} \cdot \mathbf{b}_{M,k}^{\ell-1}$  and compute  $\mathbf{b}_{R,k}^s = \mathbf{b}_{M,k}^s - w_k \mathbf{I}_s^k$ .

*Remark 5.14* (Complexity of the sampling of  $\dot{W}$  for MLQMC). The overall complexity of this sampling strategy is  $O(\mathcal{N}_S m_e + \mathcal{N}_{\mathcal{L}_\ell} \mathcal{N}_{\mathcal{L}_\ell})$  in the standard Haar wavelet case and  $O(\mathcal{N}_S m_e^\ell + \mathcal{N}_{\mathcal{L}_\ell})$  in the compactly supported case (cf. Remark 5.2), where (cf. Remark 5.8)  $\mathcal{N}_{S_\ell}$  is the number of cells of the three-way supermesh on the MLQMC level  $\ell$ ,  $m_e^\ell$  is the number of dofs of  $V^\ell$  per cell of  $D_h^\ell$  and  $N_{\mathcal{L}_\ell}$  is the number of wavelets that have non-zero support over any of the  $\mathcal{N}_{\mathcal{L}_\ell}$  cells of  $D_{\mathcal{L}_\ell}$ . Again, thanks to Theorem 4.3, we have that  $\mathcal{N}_{S_\ell} = O(\mathcal{N}_h^\ell + \mathcal{N}_{\mathcal{L}_\ell})$ , where  $\mathcal{N}_h^\ell$  is the number of cells of  $D_h^\ell$ , i.e. the number of cells of the supermesh is bounded above by a linear function of the number of cells of the parent meshes. Again, as in Remark 5.8, picking  $\mathcal{L}_\ell$  on each MLQMC level so that  $\mathcal{N}_h^\ell$  dominates yields an overall  $O(\mathcal{N}_h^\ell)$  linear cost complexity for the sampling.

*Remark 5.15* (MLQMC in the affine case). Similarly as in Remark 5.12, in the case in which the mapping to the reference cell of the FEM basis chosen is affine (e.g. Lagrange elements on simplices) a simpler, more efficient sampling strategy consists in sampling  $\dot{W}$  on the supermesh and then transfer the result onto the parent FEM meshes. The sampling of  $\dot{W}_{\mathcal{L}}$  is then the same as in the affine QMC case (see 5.12), but on the three-way supermesh. The sampling of  $\dot{W}_R$  is also very similar and relies on the affine MLMC strategy presented in Chapter 4 for point 4 of the sampling algorithm above.

*Remark 5.16* (Simpler cases: nested meshes and  $p$ -refinement). When the MLQMC mesh hierarchy is nested and/or the hierarchy is composed by taking a single mesh and increasing the polynomial degree of the FEM subspaces we have  $V^{\ell-1} \subseteq V^\ell$ . In this case everything we discussed still applies with the following simplifications: only a two-way supermesh between  $D_h^\ell$  and  $D_{\mathcal{L}^\ell}$  is needed on each MLQMC level in the  $h$ -refinement case. In the  $p$ -refinement case we only have one FEM mesh  $D_h$  and a single two-way supermesh construction is needed between  $D_h$  and the finest Haar mesh  $D_{\mathcal{L}_L}$ .

*Remark 5.17* (Non-nested mesh hierarchies and embedded domains). In practice, we assume that we are given a user-provided hierarchy  $\{G_h^\ell\}_{\ell=0}^L$  of possibly non-nested FEM meshes of the domain  $G$  on which we need the Matérn field samples. From this we construct a boxed domain  $D$  s.t.  $G \subset\subset D$  and a corresponding hierarchy of Haar meshes  $\{D_{\mathcal{L}^\ell}\}_{\ell=0}^L$  and of FEM meshes of  $D$ ,  $\{D_h^\ell\}_{\ell=0}^L$ . As in Chapter 4, it is convenient to construct each  $D_h^\ell$  so that  $G_h^\ell$  is nested within it, so that each Matérn field sample can be transferred exactly and at negligible cost on the mesh on which it is needed (this is the embedded domain strategy proposed in [132]). For simple problems, it is also possible to avoid a three-way supermesh construction by either using a nested mesh hierarchy of  $G$  and/or by solving equation (5.10) on a uniform structured mesh hierarchy so that each Haar mesh is nested within the corresponding  $D_h^\ell$ . The former simplification is typically not an option in many practical computations on complex geometries where a nested hierarchy is normally not available (cf. Section 4.2.3). The latter simplification is inadvisable when the user-provided hierarchy  $\{G_h^\ell\}_{\ell=0}^L$  is made of graded or locally refined meshes as it is desirable for the Matérn field to be sampled with the same accuracy as the solution of the PDE for which the sample is needed.

*Remark 5.18* (General wavelets). We expect the generalisation of the presented sampling methods to generic wavelets to be straight-forward, although it is unclear as to whether this would bring any considerable advantage. We leave this investigation to future research.

*Remark 5.19* (General domains). We briefly speculate on the extension of the methods presented to more general domains. The same sampling method could be generalised to general convex domains by introducing “generalised” Haar wavelets and meshes, obtained by partitioning a mesh into sub-regions and defining the cells of  $D_{\mathcal{L}}$  through aggregation of the cells of  $D_h$ . Establishing any theoretical results in this case would be more complex, but the same algorithm should carry forward after accounting for the fact that the “Haar cells” obtained through aggregation would have variable volume. The advantage of doing this is that no supermesh would then be required

in the QMC case (the Haar mesh would be nested within  $D_h$  by construction) and only one supermesh construction would be needed (between  $D_h^\ell$  and  $D_h^{\ell-1}$ ) in the non-nested MLQMC case. We leave the implementation of this extension to future work.

## 5.6 Numerical results

### 5.6.1 Testing the supermesh construction

As mentioned in remarks 5.8 and 5.14, the sampling strategies presented have linear cost complexity in the number of supermesh cells. In turn this number is proportional to the number of cells of each of the parent meshes. Before testing the sampling algorithms just presented, we therefore use the libsupermesh [121] library to get an idea of how large the proportionality constant can be in practice.

Unfortunately there is no dedicated three-way supermeshing algorithm in libsupermesh and we therefore first compute a two-way supermesh between the FEM meshes and then another two-way supermesh between this “FEM” supermesh and the Haar mesh. We recall that since the Haar mesh is structured knowing for each “FEM” supermesh cell which Haar cells it intersects is inexpensive (cf. Remark 5.7). However, our approach is likely to produce three-way supermeshes with more cells than actually needed (the optimal strategy would be to mesh the intersection of cells from all the three meshes involved at the same time).

We take the same non-nested unstructured mesh hierarchies of the domain  $D = (-1, 1)^d$  that we used in Chapter 4 and we count the number of cells resulting from the constructions needed for both the QMC and MLQMC sampling strategies. Results are shown in Tables 5.1 and 5.2 for the 2D and 3D case respectively. We gradually increase the Haar level with the MLQMC level so that the Haar meshes have cell size comparable to the finer of the two FEM meshes on each MLQMC level. In 2D, the ratio  $\mathfrak{R}^\ell = \mathcal{N}_{S_\ell}/(\mathcal{N}_h^\ell + \mathcal{N}_{\mathcal{L}_\ell})$  plateaus to a constant value in both the QMC and MLQMC case and no significant difference is seen when increasing the Haar levels of the whole hierarchy.

In 3D the behaviour as the MLQMC level increases looks similar, although it is harder to be sure given the smaller hierarchy. What is different though is that the ratio  $\mathfrak{R}^\ell$  seems to increase as the Haar level increase. We therefore investigate this behaviour further by studying how the ratios  $\mathfrak{R}^\ell$  and  $\mathcal{N}_{S_\ell}/\mathcal{N}_{\mathcal{L}_\ell}$  behave as the Haar level  $\mathcal{L}_\ell$  increases for a fixed pair of FEM meshes (the first two in the unstructured hierarchy). Results are shown in Figure 5.4. Both ratios appear to be asymptotically bounded as  $\mathcal{L}_\ell$  grows showing that  $\mathcal{N}_{S_\ell} = O(\mathcal{N}_h^\ell + \mathcal{N}_{\mathcal{L}_\ell})$  as stated by Theorem 4.3. Both Table 5.2 and Figure 5.4 suggest, however, that in 3D the proportionality constant is fairly large. Whether the sub-optimal supermeshing approach used (see above) is to be blamed for such a large constant is unclear and we leave this to future research.

In Tables 5.1 and 5.2 we report the number of supermesh cells of MLQMC hierarchy in which the Haar level is increased as the FEM mesh is refined so that Haar

cells and FEM cells have comparable sizes. As we see further on in this section, in the MLQMC case there is little benefit in using high Haar levels on fine MLQMC levels and the Haar level can be kept low on fine MLQMC levels. If this is the case, the effect of including the Haar mesh in the supermesh construction is negligible and the number of cells in the three-way supermesh is comparable to the number of cells of the two-way supermesh between the FEM meshes only (cf. Tables 4.1 and 4.2 in Section 4.5).

				$\mathfrak{R}^\ell (\mathcal{L}_\ell = \ell)$	$\mathfrak{R}^\ell (\mathcal{L}_\ell = \ell + 1)$		
$\ell$ (2D)	$h_\ell$	$ \square_{\mathcal{L}_\ell} ^{1/2}$	$n_\ell$	QMC	MLQMC	QMC	MLQMC
1	0.707	1	32	0.67	n/a	1.33	n/a
2	0.416	0.5	120	1.63	2.26	2.47	2.88
3	0.194	0.25	500	2.69	4.38	3.41	4.84
4	0.098	0.125	2106	3.41	5.82	3.96	6.02
5	0.049	0.0625	8468	3.77	6.44	4.24	6.52
6	0.024	0.03125	33686	3.95	6.77	4.36	6.75
7	0.012	0.01563	134170	4.05	6.93	4.43	6.85
8	0.006	0.00781	535350	4.09	7.01	4.46	6.91
9	0.003	0.00391	2143162	4.11	7.05	4.47	6.94

Table 5.1: Testing a 2D unstructured non-nested hierarchy over  $(-1, 1)^2$ : mesh level  $\ell$ , maximal cell size  $h_\ell$ , Haar cell size  $|\square_{\mathcal{L}_\ell}|^{1/2}$  for  $\mathcal{L}_\ell = \ell$ , number of FEM mesh cells  $n_\ell$ , ratio  $\mathfrak{R}^\ell = \mathcal{N}_{S_\ell}/(\mathcal{N}_h^\ell + \mathcal{N}_{\mathcal{L}_\ell})$ . Note that the FEM mesh cell size roughly decreases proportional to  $2^{-\ell}$  and the ratio  $\mathfrak{R}^\ell$  seems to plateau asymptotically as  $\ell$  grows.

				$\mathfrak{R}^\ell (\mathcal{L}_\ell = \ell - 1)$	$\mathfrak{R}^\ell (\mathcal{L}_\ell = \ell)$		
$\ell$ (3D)	$h_\ell$	$ \square_{\mathcal{L}_\ell} ^{1/3}$	$n_\ell$	QMC	MLQMC	QMC	MLQMC
1	0.866	1	384	0.98	n/a	0.86	n/a
2	0.437	0.5	7141	4.09	16.8	19.7	53.6
3	0.280	0.25	22616	10.5	195	43.3	413
4	0.138	0.125	190081	13.3	153	48.2	326
5	0.070	0.0625	1519884	14.8	173	51.0	358

Table 5.2: Testing a 3D unstructured non-nested hierarchy over  $(-1, 1)^3$  (part 1): mesh level  $\ell$ , maximal cell size  $h_\ell$ , Haar cell size  $|\square_{\mathcal{L}_\ell}|^{1/3}$  for  $\mathcal{L}_\ell = \ell$ , number of FEM mesh cells  $n_\ell$ , ratio  $\mathfrak{R}^\ell = \mathcal{N}_{S_\ell}/(\mathcal{N}_h^\ell + \mathcal{N}_{\mathcal{L}_\ell})$ . Note that the FEM mesh cell size roughly decreases proportional to  $2^{-\ell}$ .

In Tables 5.3 and 5.4 we present some representative average CPU timings over multiple ( $N = 1000$  in 2D and  $N = 100$  in 3D) realisations of white noise (both  $\dot{\mathbf{W}}_{\mathcal{L}}$  and  $\dot{\mathbf{W}}_R$  terms) and Matérn field samples in 2D ( $\mathcal{L} = \ell + 1$ ) and 3D ( $\mathcal{L} = \ell$  and  $\mathcal{L} = 4$ ) over  $(-1, 1)^d$  and with  $\nu = 2 - d/2$ . These timings have been obtained

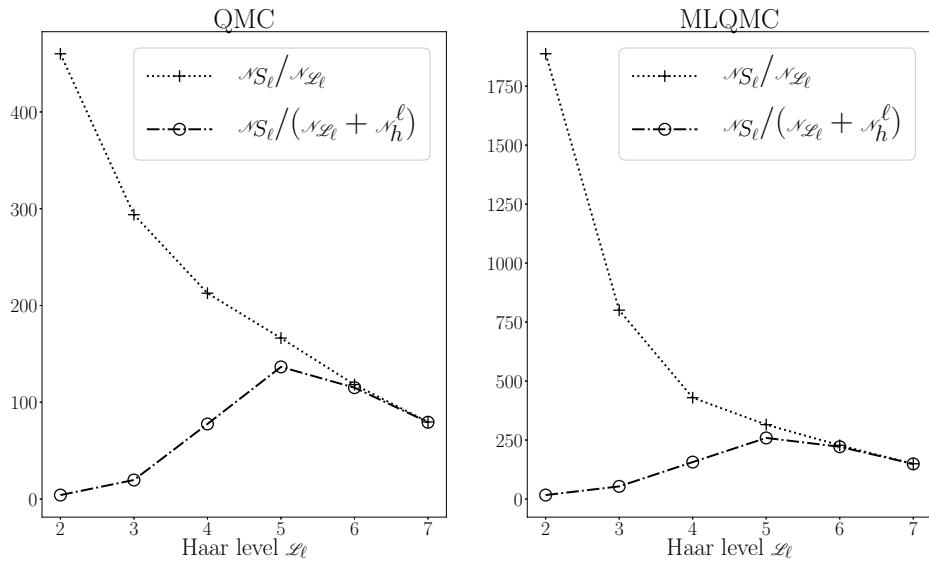


Figure 5.4: Testing a 3D unstructured non-nested hierarchy over  $(-1, 1)^3$  (part 2): we fix the FEM meshes with  $\ell = 1, 2$  in Table 5.2 and we increase the Haar level  $\mathcal{L}$ .  $\mathcal{N}_S$  is the number of supermesh cells. In the QMC case (left) this is a two-way supermesh between the Haar mesh and the FEM mesh for  $\ell = 2$ ; in the MLQMC case (right), this is a three-way supermesh between the Haar mesh and the FEM meshes for  $\ell = 1, 2$ .  $\mathcal{N}_{\mathcal{L}}$  is the number of Haar cells and  $\mathcal{N}_h$  is the number of cells of the FEM mesh with  $\ell = 2$ . Both ratios appear to be bounded as the Haar level  $\mathcal{L}$  grows. The ratio  $\mathcal{N}_S/(\mathcal{N}_h + \mathcal{N}_{\mathcal{L}})$  (dashed line) initially increases with the Haar level before decreasing again, mirroring the behaviour observed for  $\mathfrak{R}^\ell$  in the last four columns of Table 5.2.

in serial by using a single thread on a computing node with an Intel® Xeon® Gold 6140 CPU (2.30GHz). We consider both the QMC and MLQMC case. For the latter, the timings correspond to the cost of sampling coupled realisations of white noise and Matérn field samples on MLQMC levels  $\ell$  and  $\ell - 1$ . These timings do not include offline one-off costs such as the assembly of the coupled linear systems deriving from the discretisation of Equations (5.73) and (5.74). We note that both in 2D and in 3D the cost of sampling white noise realisations is always larger than the cost of solving the coupled linear system. However, there is most likely room for improvement considering that: 1) our implementation of the white noise sampling routine is currently Python-based (while the solver uses PETSc routines which are heavily optimised). 2) The three-way supermeshing algorithm used is suboptimal. We expect that in the 2D case it should be possible to compute white noise samples faster than solving (5.73) and (5.74). In the 3D MLQMC case, capping the maximum Haar level could bring further improvements since this does not seem to affect MLQMC convergence (see next subsection). Overall, these timings could largely benefit from a more efficient implementation, which we leave to future work.

*Remark 5.20.* The QMC and MLQMC methods presented are designed to work with non-nested unstructured grid hierarchies. In this setting, a supermesh construction might be unavoidable even in the case in which the field is sampled via an alternative

method such as e.g. circulant embedding [49], since this is the only way in which the sampled field can be integrated exactly on the target non-nested unstructured grid. For low smoothness fields ( $\nu \leq 1$ ), Graham et al. have proved that there is no loss in the convergence rate due to non-nested interpolation [74]. However, this operation, albeit faster, still introduces extra bias, and it is likely to still harm convergence when smoother fields are used. All in all, we remark that the high white noise sampling costs observed in the QMC case stem directly from the supermesh construction, and that similar timings would also be observed if a supermesh construction were used to integrate a field sampled on a structured grid over a target non-nested mesh. From this point of view, the MLQMC sampling costs are “only” a factor of 10 more expensive than for QMC. Finally, in our (ML)QMC sampling strategy, the white noise term is always sampled exactly (on the FEM space) independently on the Haar level choice. This might enable the use of lower Haar levels than for algorithms that sample white noise inexactly since for our strategy we are not restricted by bias.

$\ell$	2	3	4	5	6	7	8	9
W sample (QMC)	$1 \times 10^{-3}$	$3 \times 10^{-3}$	$8 \times 10^{-3}$	0.030	0.13	0.58	2.54	10.9
W sample (MLQMC)	$1 \times 10^{-3}$	$5 \times 10^{-3}$	0.020	0.089	0.38	1.36	5.84	23.9
solver	$6 \times 10^{-4}$	$9 \times 10^{-4}$	$3 \times 10^{-3}$	0.010	0.05	0.21	0.95	3.98

Table 5.3: Average timings over 1000 realisations for the coupled sampling of coupled white noise and Matérn fields ( $\nu = 1$ ) for QMC and MLQMC in 2D. The timings do not include offline one-off costs, such as the set-up of the linear solver. The meshes used are the same as in Table 5.1 and we use  $\mathcal{L}_\ell = \ell + 1$ . We note that the cost of sampling white noise is larger than the cost of solving (5.73) and (5.74) by a factor of 3 for QMC and of 6 for MLQMC.

$\ell$	1	2	3	4	5
W sample (QMC)	$2 \times 10^{-3}$	0.04	0.52	4.89	40.8
W sample (MLQMC)	$2 \times 10^{-3}$	0.16	5.51	61	495
W sample (MLQMC) $\mathcal{L} = 4$	0.051	0.71	5.51	21	125
solver	$4 \times 10^{-4}$	$4 \times 10^{-3}$	0.03	0.23	1.85

Table 5.4: Average timings over 100 realisations for the coupled sampling of coupled white noise and Matérn fields ( $\nu = 1/2$ ) for QMC and MLQMC in 3D. The timings do not include offline one-off costs, such as the set-up of the linear solver. The meshes used are the same as in Table 5.2 and we use  $\mathcal{L}_\ell = \ell$  and  $\mathcal{L} = 4$  (only where indicated). We note that the cost of sampling white noise is much larger than the cost of solving (5.73) and (5.74). This stems from the large number of cells of the supermeshes involved, cf. Table 5.2. The cost on the fine levels can be decreased by capping the Haar level on the finer MLQMC levels since this does not seem to affect MLQMC convergence (but it does affect QMC, see next subsection).

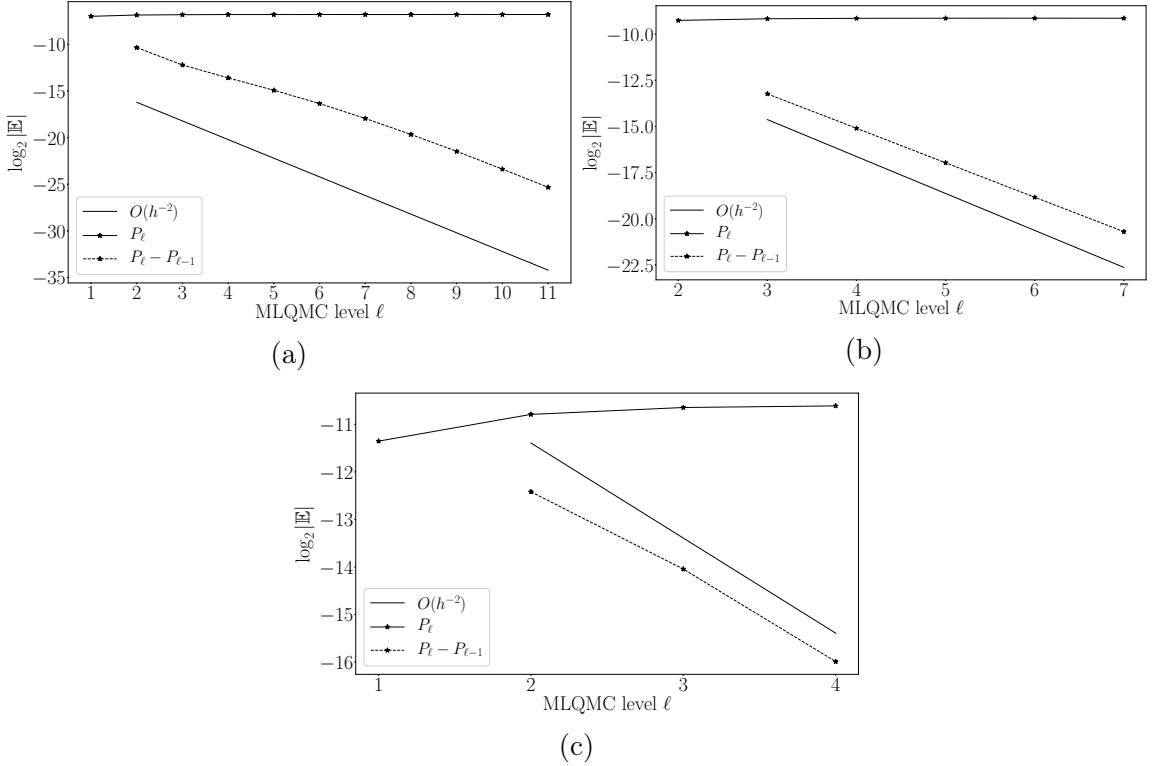


Figure 5.5: Logarithm of the absolute value of the expected value of  $P_\ell$  and  $P_\ell - P_{\ell-1}$  as a function of the MLQMC level  $\ell$  in (a) 1D, (b) 2D and (c) 3D. We observe a decay rate of  $O(h^2)$  in all dimensions. These results are independent from the Haar level chosen as we always compute the exact action of white noise independently from the choice of  $\mathcal{L}$ .

### 5.6.2 Testing the sampling strategies

We now test our sampling algorithms. We consider test problem (5.49) over the domain  $G = (-0.5, 0.5)^d$  with forcing term  $f = 1$ , i.e. we solve

$$\begin{aligned} -\nabla \cdot (e^{u(\mathbf{x}, \omega)} \nabla p(\mathbf{x}, \omega)) &= 1, & \mathbf{x} \in G = (-0.5, 0.5)^d, & \omega \in \Omega, \\ p(\mathbf{x}, \omega) &= 0, & \mathbf{x} \in \partial G, & \omega \in \Omega, \end{aligned} \quad (5.81)$$

where  $u(\mathbf{x}, \omega)$  is a Matérn field sampled by solving equation (5.10) over  $D = (-1, 1)^d$  with  $\lambda = 0.25$  and mean and standard deviation chosen so that  $\mathbb{E}[e^u] = 1$ ,  $\mathbb{V}[e^u] = 0.2$ . For simplicity, we take  $P(\omega) = \|p\|_{L^2(G)}^2(\omega)$  as our output functional of interest.

As in Chapter 3, we solve equations (5.10) and (5.81) with the FEniCS software package [120]. For simplicity, we stick to the  $h$ -refinement case and we discretise the equations using continuous piecewise linear Lagrange elements. We employ the conjugate gradient routine of PETSc [15] preconditioned by the BoomerAMG algebraic multigrid algorithm from Hypre [57] for the linear solver for both equations. Again, we declare convergence when the absolute size of the preconditioned residual norm is below a tolerance of  $10^{-10}$ . We still use the libsupermesh software package [121] for the supermesh constructions. We use random digital shifted Sobol' sequences sampled

with a custom-built Python and C wrapper of the Intel® Math Kernel Library Sobol' sequence implementation augmented with Joe and Kuo's primitive polynomials and direction numbers [98] (maximum dimension = 21201).

We construct the mesh hierarchies  $\{G_h^\ell\}_{\ell=0}^L$  and  $\{D_h^\ell\}_{\ell=0}^L$  so that, for all MLQMC levels  $\ell$ ,  $G_h^\ell$  is nested within  $D_h^\ell$ , yet  $G_h^{\ell-1}$  and  $D_h^{\ell-1}$  are not nested within  $G_h^\ell$  and  $D_h^\ell$  respectively. Even if we do not exploit it anywhere in our implementation, we take all the meshes in both hierarchies to be simplicial, uniform and structured for simplicity with mesh sizes  $h_\ell = 2^{-(\ell+1)}$  in 1D,  $h_\ell = 2^{-1/2} 2^{-\ell}$  in 2D and  $h_\ell = \sqrt{3} 2^{-(\ell+1)}$  in 3D.

### 5.6.3 MLQMC convergence

Unlike in Section 4.5, we do not consider functionals of the Matérn field  $u$  and we directly focus on the estimation of  $\mathbb{E}[P]$ . The reason is that in 2D and 3D the smoothness of  $u$  is low and we only observe standard Monte Carlo convergence rates in numerical experiments (not shown).

We first study how the quantities  $|\mathbb{E}[P_\ell]|$  and  $|\mathbb{E}[P_\ell - P_{\ell-1}]|$  vary as the MLQMC level is increased. As in Section 4.5, assuming  $u$  can be sampled exactly, we expect the MLMC parameter value  $\alpha$  to be  $\alpha = \min(\nu + 1, p + 1)$  [80]. Numerical results are shown in Figure 5.5 where observe a decay rate of  $\alpha = 2$  in 1D, 2D and 3D. In 3D we might have expected the rate to be 1.5 due to the lack of smoothness of the coefficient  $u$  which is only in  $C^{0.5-\epsilon}(\bar{G})$  for any  $\epsilon > 0$  [80]. However, at the discrete level the FEM approximation of  $u$  is in  $C^{1,\epsilon}(G)$  a.s. for any  $\epsilon > 0$  and we might be observing a pre-asymptotic regime.

As a next step, we analyse the convergence behaviour of QMC and MLQMC with respect to the number of samples. In Theorem 5.6 we showed that in the QMC case we expect an initial QMC-like convergence rate followed by a standard MC rate of  $O(N^{-1/2})$  and that the higher the Haar level is, the later the transition between the two regimes happens. No results regarding the MLQMC case were derived, but we expect a similar behaviour to occur. Furthermore, we would like to determine whether the multilevel technique can improve on QMC by bringing further variance reduction.

We draw inspiration from the original MLQMC paper by Giles and Waterhouse [67] and we study the convergence behaviour of both QMC and MLQMC as the MLQMC level is increased. Results are shown in Figure 5.6. We increase the Haar level with the MLQMC level so that the Haar mesh size is always proportional to the FEM mesh size, but we consider two different strategies: 1) we choose the Haar mesh size to be comparable to the FEM mesh size (Figures 5.6a, 5.6c, 5.6e) and 2) we pick the Haar mesh size to be smaller than the FEM mesh size (Figures 5.6b, 5.6d, 5.6f). For both scenarios, we compute the variance  $\mathbb{V}_\ell$  of the (ML)QMC estimator on MLQMC level  $\ell$  by using  $M = 128$  ( $M = 64$  in 3D) randomisations of the Sobol' sequence used and we monitor the quantity  $\log_2(N\mathbb{V}_\ell)$  as the number of samples  $N$  is increased. Various colours are used in Figure 5.6 to indicate the different sample sizes. The horizontal lines correspond to QMC and the oblique lines to MLQMC.

For standard Monte Carlo and MLMC, we have  $\mathbb{V}_\ell = O(N^{-1})$ , giving a  $\log_2(N\mathbb{V}_\ell)$  of  $O(1)$ . For this reason, if we were observing a MC-like convergence rate, we would see the different coloured lines of Figure 5.6 overlapping. The fact that this does not

happen means that we are in fact observing a QMC-like rate which is faster than  $O(N^{-1})$  (for the variance). However, it is clear by looking at Figures 5.6a, 5.6c and 5.6e that as  $N$  grows the lines start overlapping, marking a decay to a  $O(N^{-1})$  rate of convergence (for the variance) as predicted by Theorem 5.6. By comparing the figures on the left hand side to those on the right-hand side, it is also clear that increasing the Haar level delays the occurrence of this behaviour both in the QMC case (as predicted by Theorem 5.6) and in the MLQMC case. Furthermore, it appears that in the MLQMC case the convergence rate decays sooner than in the QMC case. Finally, we note that MLQMC indeed benefits from the combination of QMC and MLMC: the variance of the MLQMC estimator on any level is always smaller than the corresponding QMC estimator for the same number of samples, with large variance reductions on the fine levels.

In Figure 5.7 we take a closer look at the 2D examples from Figure 5.6 and we compare the observed (ML)QMC convergence rate with the rate that would be expected in a standard MC regime. We monitor the standard deviation of the (ML)QMC estimators for  $P_\ell$  (left figures) and  $P_\ell - P_{\ell-1}$  (right figures). Initially, a QMC-like convergence rate is observed for both quantities, which eventually decays to a standard MC rate. This transition occurs later when the Haar level is larger. To see this, compare the top (lower Haar level) and bottom (higher Haar level) plots and the different colour lines (in this test case a higher (ML)QMC level corresponds to a higher Haar level) in Figure 5.7.

We now focus on the 2D case only for simplicity and see how both QMC and MLQMC perform in practice when applied to equation (5.81). In Figure 5.8 we study the adaptivity and cost of (ML)QMC as the root mean square error tolerance  $\varepsilon$  is decreased for the same FEM hierarchy, but for two different Haar level hierarchies. The top plots in the figure correspond to the choice of Haar meshes with mesh size comparable to the FEM mesh size ( $|\square_{\mathcal{L}_\ell}|^{1/2} = 2^{-\ell}$ ). The results in the bottom plots are instead obtained by fixing the Haar level to  $\mathcal{L}_\ell = 6$  for all  $\ell$ . In both cases we fix the number of randomisations to be  $M = 32$ .

In the plots on the left hand side in Figure 5.8 we see how MLQMC automatically selects the number of samples according to the greedy strategy highlighted in Section 2.2 so as to satisfy the given error tolerance. As in the MLMC case, more samples are taken on coarse levels and only a few on the fine levels. The second Haar level strategy (plot on the bottom right) uses higher Haar levels on the coarse MLQMC levels, which corresponds to a later decay to a  $O(N^{-1/2})$  rate (cf. Figure 5.6). Therefore this strategy requires lower sample sizes (compare with the top left plot).

In the plots on the right-hand side we show the overall cost of QMC and MLQMC as the root mean square error tolerance  $\varepsilon$  is reduced. More specifically, we plot the quantity  $\varepsilon^2 C_{\text{tot}}$ , where  $C_{\text{tot}}$  is the total cost. The reason is that the total cost complexity of MLMC for this problem (MLMC parameters:  $\beta = 2\alpha = 4$ ,  $\gamma = 2$ , cf. Section 2.2) is  $O(\varepsilon^{-2})$ , giving the  $\varepsilon^2 C_{\text{tot}}$  factor to be  $O(1)$  for all  $\varepsilon$ . The fact that the MLQMC cost line is not horizontal, but decreases as  $\varepsilon$  is reduced shows that the total complexity of MLQMC is better than  $\varepsilon^{-2}$ , i.e. that our MLQMC algorithm has a better-than-MLMC complexity. This improved complexity stems from the fact that we are observing a QMC-like convergence rate with respect to  $N$ .

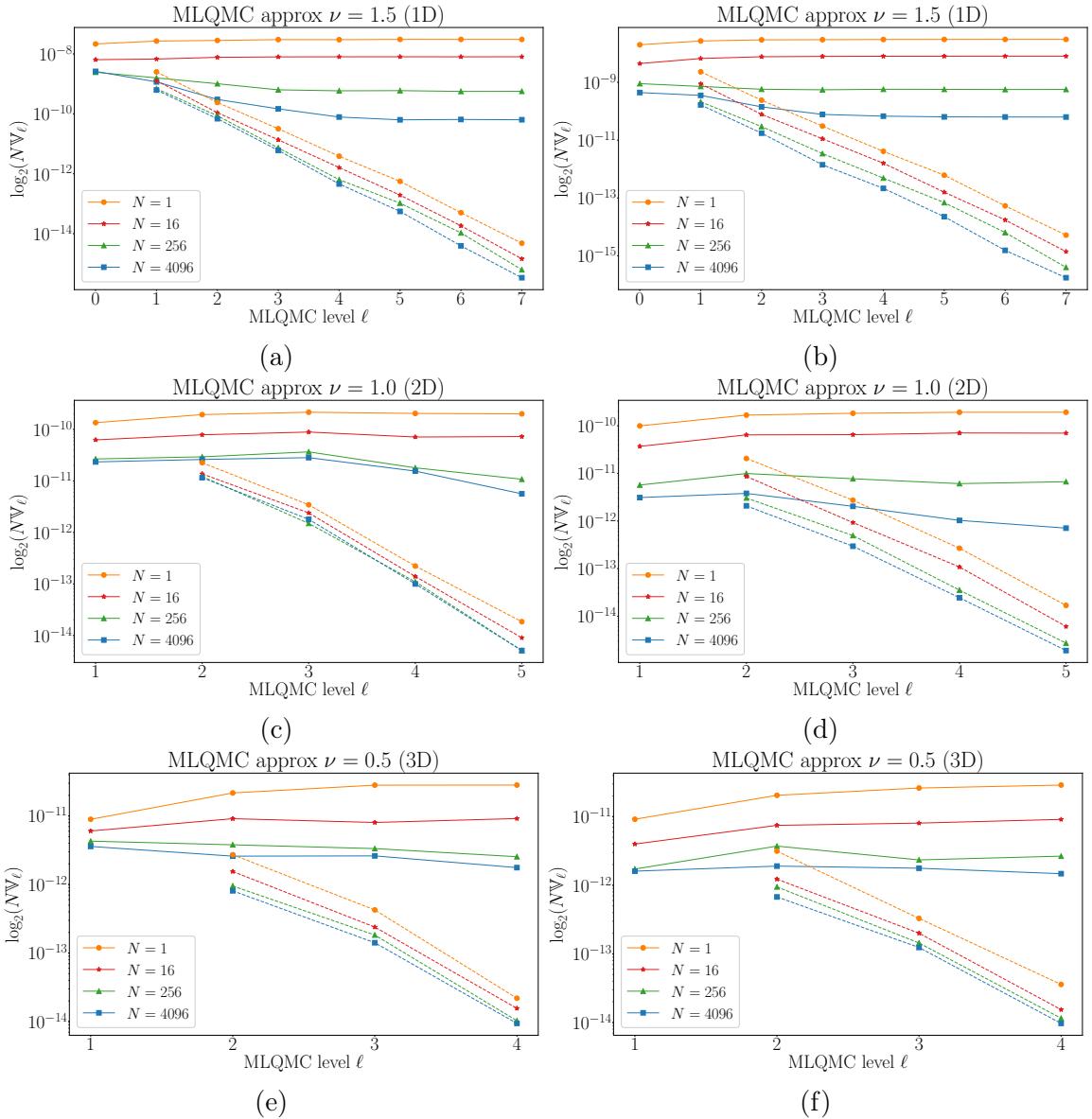


Figure 5.6: Convergence behaviour of (ML)QMC with respect to the number of samples  $N$  in 1D (a)-(b) and 2D (c)-(d) ( $M = 128$ ), and in 3D (e)-(f) ( $M = 64$ ). Plots (a),(c),(e) (on the left) are obtained by choosing Haar mesh sizes comparable to the FEM mesh sizes:  $|\square_{\mathcal{L}_\ell}| = 2^{-(\ell+1)}$  in 1D (a), and  $|\square_{\mathcal{L}_\ell}|^{1/d} = 2^{-\ell}$  in 2D (c) and 3D (e). Plots (b),(d),(f) (on the right) are obtained by choosing Haar meshes which are finer than the corresponding FEM meshes:  $|\square_{\mathcal{L}_\ell}| = 2^{-(\ell+2)}$  in 1D (b),  $|\square_{\mathcal{L}_\ell}|^{1/2} = 2^{-(\ell+2)}$  in 2D (d) and  $|\square_{\mathcal{L}_\ell}|^{1/3} = 2^{-(\ell+1)}$  in 3D (f). The (approximately) horizontal and oblique lines correspond to QMC and MLQMC respectively. Different colours indicate different sample sizes. On the y-axis we monitor (the logarithm of) the product between  $N$  and the (ML)QMC estimator variance on level  $\ell$ . This product is  $O(1)$  when the convergence rate is MC-like and therefore the coloured lines overlap if a  $O(N^{-1/2})$  MC rate is observed. In the figure we observe a pre-asymptotic QMC-like convergence rate that then tails off to a standard MC rate (the lines initially do not overlap, but they start overlapping as  $N$  is increased). This phenomenon always occurs (cf. Theorem 5.6), but it happens later when the Haar level is increased (figures on the right).

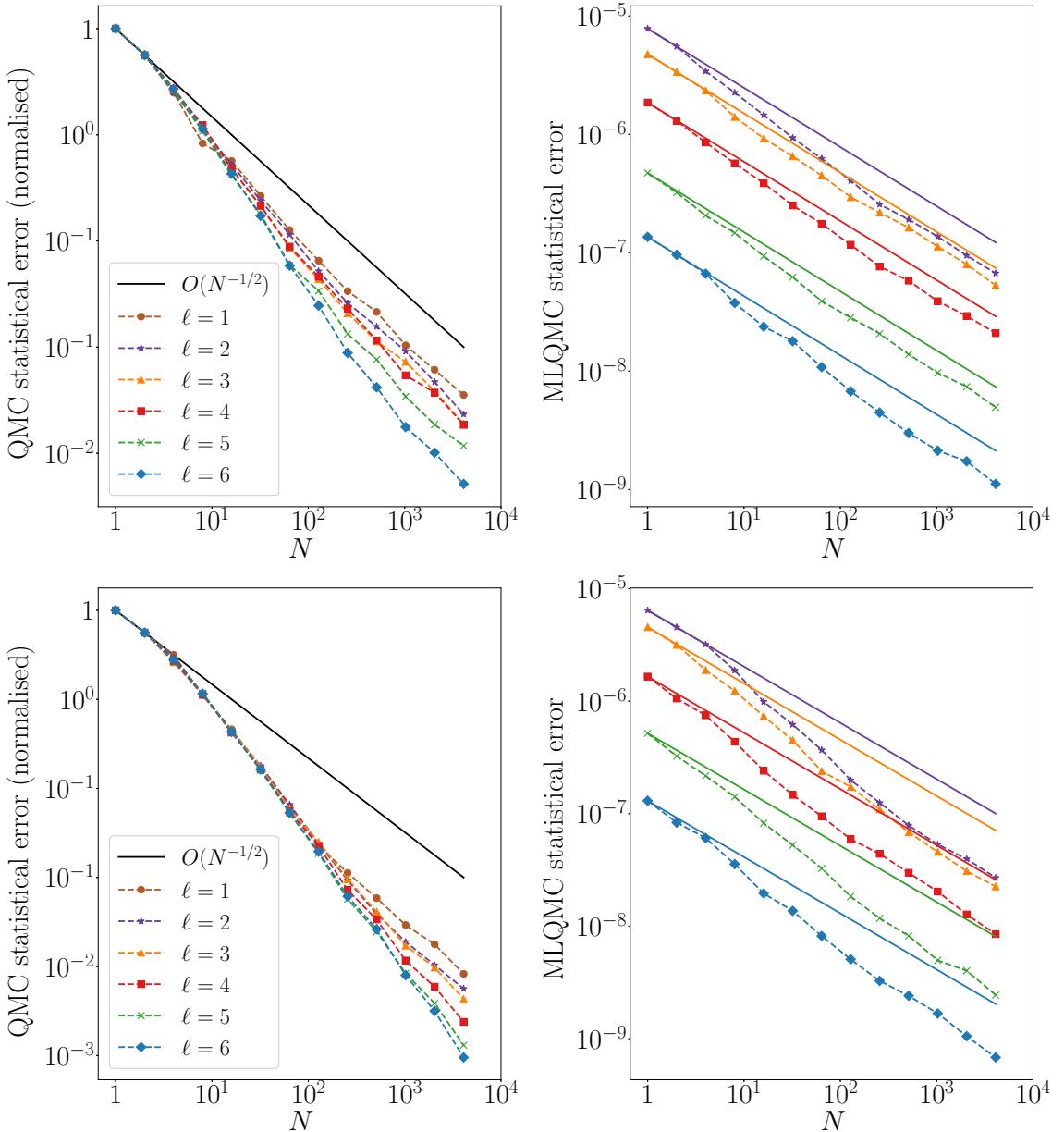


Figure 5.7: Convergence behaviour of (ML)QMC with respect to the number of samples  $N$  in 2D. The top and the bottom plots correspond to the same problem and (ML)QMC set-up considered in Figures 5.6c and 5.6d respectively. The continuous lines correspond to the standard MC rate. In the figures on the left, we plot the standard deviation (SD) of the QMC estimator for  $P_\ell$  as a function of the number of samples. In order to make the figures clearer, we normalise this value by dividing by the SD of the estimator when  $N = 1$ . The actual values can be recovered by looking at the horizontal orange lines in Figures 5.6c and 5.6d. In the figures on the right, we plot the SD of the MLQMC estimator for  $P_\ell - P_{\ell-1}$  as a function of  $N$  (dashed lines) and we compare them with the theoretical standard MC convergence behaviour (continuous lines). Overall, we observe a pre-asymptotic QMC rate that eventually decays to a  $O(N^{-1/2})$  standard MC rate. By comparing the different figures and the behaviour on different levels (here larger (ML)QMC levels correspond to larger Haar levels), we note how this decay occurs later for larger Haar levels, as expected (cf. Theorem 5.6).

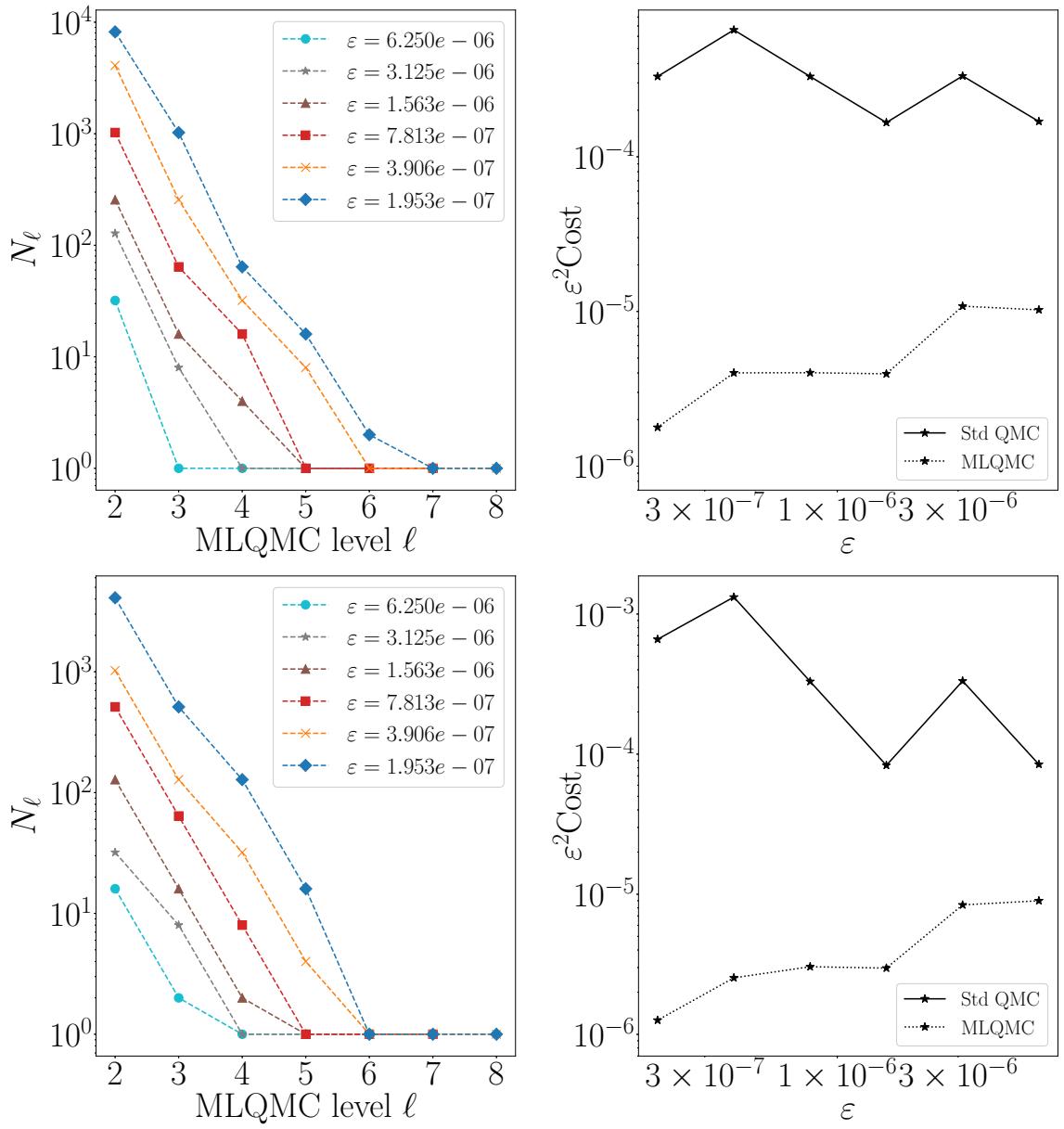


Figure 5.8: *MLQMC convergence for the solution of (5.81)*. We take  $M = 32$  and consider two Haar level hierarchies:  $\mathcal{L}_\ell = 2 + \ell$  (top plots) and  $\mathcal{L}_\ell = 6$  for all  $\ell$  (bottom plots). In the plots on the left we show how the MLQMC algorithm automatically selects the optimal number of samples  $N_\ell$  on each level to achieve a given tolerance  $\varepsilon$ . Note that we have dropped the first mesh of the hierarchy as it is too coarse and it would not bring any significant advantage to the performance of MLQMC (same reasoning as for MLMC, see [68]). We observe that on the finest levels only one sample is used, making MLQMC equivalent to plain MLMC on these levels. In the plot on the right we compare the efficiency of MLQMC with QMC for different tolerances. MLQMC appears to have a better-than- $O(\varepsilon^{-2})$  total cost complexity and significantly outperforms QMC.

As  $\varepsilon$  is decreased, we expect the cost complexity to decay to an  $\varepsilon^{-2}$  rate: for extremely fine tolerances very large sample sizes are required yielding the asymptotic  $O(N^{-1/2})$  standard MC rate and harming the overall cost complexity. However, even in this case, the overall MLQMC cost benefits from the pre-asymptotic regime and MLQMC still outperforms MLMC (see Figure 5.9). Similarly, QMC initially benefits from a faster convergence rate with respect to  $N$ . As the tolerance is decreased, the QMC rate decays to a standard MC rate and the total cost of QMC starts increasing faster than  $O(\varepsilon^{-2})$ .

Comparing the costs between the top and bottom of Figure 5.8, it appears that increasing the Haar level on the coarse MLQMC levels improves the total MLQMC cost while, in the QMC case, decreasing the Haar level harms convergence. This suggests that the Haar level choice has a considerable impact on the overall MLQMC and QMC performance. We investigate this in Figure 5.9, where we show the total cost of (ML)QMC for different Haar level hierarchies and we compare it with the cost of standard MLMC. The  $x$ -axis and the black lines in both plots are the same. We present the costs of two versions of MLMC: the black dash-dotted line corresponds to standard MLMC, while the black dashed line corresponds to a MLMC algorithm in which the number of samples are restricted to be in powers of 2. Since this restriction also applies to MLQMC, we believe the comparison between MLMC and MLQMC to be fairer in this case, but we present both lines for completeness. We note that MLQMC outperforms MLMC by a factor of approximately 8, depending on the Haar level choice.

*Remark 5.21.* The results shown in Figures 5.8 and 5.9 do not account for differences in the cost per sample due to variations in the number of supermesh cells. If the cost of solving the PDE with random coefficients of interest (e.g. equation (5.81)) dominates over the cost of sampling white noise realisations, these results are still valid as is. Otherwise, extra care must be taken when using Haar meshes which are much finer than the corresponding FEM meshes since this results in a large number of supermesh cells. In the figures this would apply to Haar levels greater than  $\mathcal{L}_\ell = \{4, \dots, 10\}$  (gray line in the plot on the right) and there is clearly a trade-off since large  $\mathcal{L}$  means faster decay with respect to  $N$ , but larger costs per sample as well.

Recall the convergence results with respect to the number of samples shown in Figure 5.6: even if the convergence rate decreases as  $N_\ell$  increases, it is clear from Figure 5.8 (left) that this only happens on the coarse levels where more samples are needed. Since for problem (5.81) and the FEM discretisation chosen we are in the “good” case of the MLMC theorem (i.e.  $\beta > \gamma$ , cf. Theorem 2.1), the multilevel cost is dominated by the sample cost on the coarse levels. We therefore expect to obtain computational gains by increasing the Haar level on the coarse MLQMC levels. At the same time, we do not expect to lose in computational efficiency if we decrease the Haar level on the fine levels as these are not dominating the total cost. Note that in the QMC case there is only one level and the only possible strategy is to keep the Haar level as high as required.

By looking at Figure 5.9 it is clear that our expectations are met. In the QMC case (plot on the left) we see that a small Haar level results in significant cost increase for

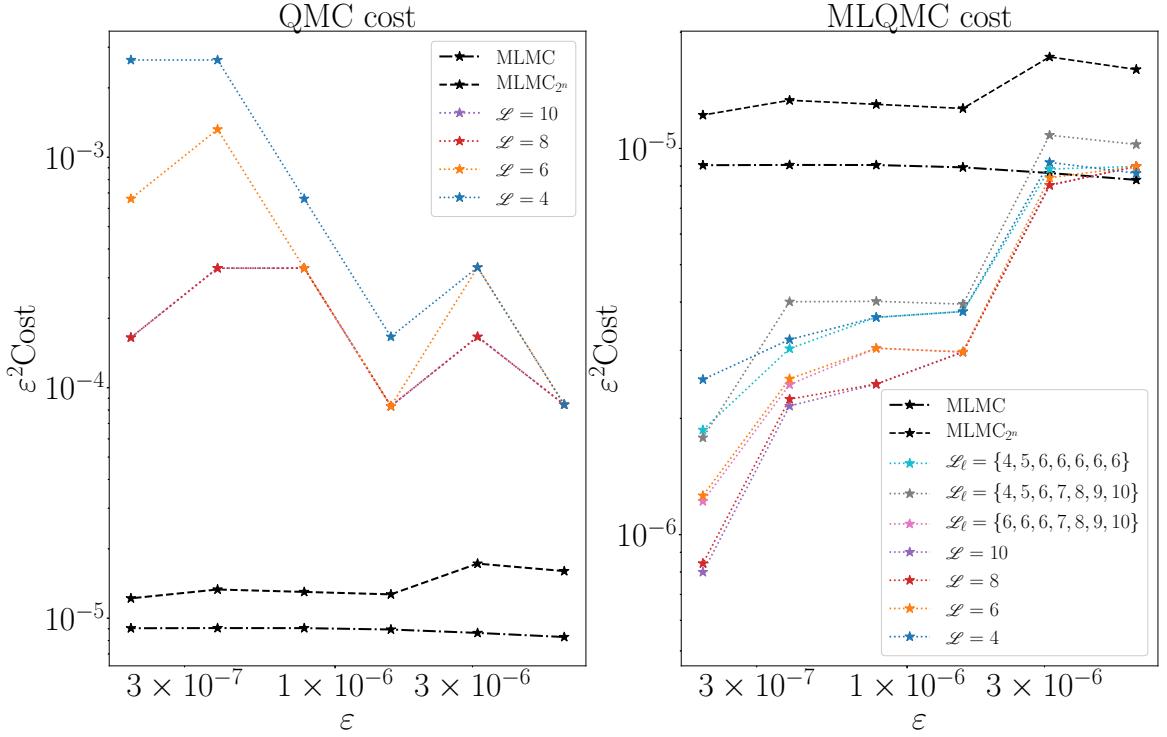


Figure 5.9: MLMC, QMC and MLQMC total computational cost needed for the solution of (5.81) with the same FEM mesh hierarchy as in Figure 5.8. In the (ML)QMC case, we take  $M = 32$  and consider different Haar level hierarchies which correspond to different computational costs. The x-axis and the MLMC lines are the same in both plots. MLQMC outperforms MLMC which in turn outperforms QMC.

small tolerances, while for large Haar levels we retain good convergence with respect to  $N$  and a cost complexity which looks just slightly worse than  $O(\varepsilon^{-2})$ . For the tolerances considered, there seems to be little advantage in increasing the Haar level beyond the  $\mathcal{L} = 8$  threshold. For this specific problem, the optimal strategy would be to increase the Haar level as the mesh is refined and set  $\mathcal{L}_\ell = \{4, 5, 6, 6, 8, 8\}$  so that the Haar level is increased only when needed. Generally speaking, we believe that it is never advantageous to use Haar meshes much finer than the FEM meshes (cf. Remark 5.21).

In the MLQMC case (plot on the right in Figure 5.9), we note that increasing the Haar level on the coarse levels indeed brings computational advantages (e.g. compare the gray and pink lines) and decreasing it on the fine levels does not seem to affect the total cost (e.g. compare the pink with the orange line), as predicted. The optimal strategy therefore consists of increasing the Haar level on the coarse levels and either keeping it constant across the MLQMC hierarchy or possibly even decreasing it (there is little computational advantage in decreasing it if the FEM meshes on the fine levels are already much finer than the Haar mesh). For the MLQMC hierarchy, a good choice seems to fix  $\mathcal{L}_\ell = 6$  for all  $\ell$  since a larger Haar level would significantly increase the number of supermesh cells (cf. Remark 5.21).

Overall, our MLQMC strategy outperforms MLMC, which in turn outperforms

QMC. Standard MC is always worse than QMC, by up to two orders of magnitude for small  $\varepsilon$  (not shown).

*Remark 5.22.* The optimal Haar strategy is likely to change if the problem to be solved belongs to the other two cases of the MLMC theorem (Theorem 2.1), i.e.  $\beta = \gamma$  or  $\beta < \gamma$ . In the former case ( $\beta = \gamma$ ), the total multilevel cost is simultaneously dominated by all levels in the multilevel hierarchy and we believe in this case that the optimal strategy is to use a Haar mesh hierarchy of mesh sizes comparable to the FEM mesh sizes (e.g. as for the gray line in the right plot of Figure 5.9). In the latter case ( $\beta < \gamma$ ), the total multilevel cost is dominated by the fine levels. In this case it might be advantageous to keep the Haar level low on the coarse levels and to increase it on the fine levels. We remark that all these considerations are qualitative and to obtain more precise statements on the exact Haar level choice a stronger MLQMC convergence theory than that presented in Section 5.4 would be required. However, establishing such a theory is complicated (the low smoothness of some of the Matérn fields considered is even outside of the theory presented by Hermann and Schwab [88]) and we leave this to future research.

*Remark 5.23.* If we are in the  $\beta > \gamma$  case, then the Haar level is capped on the fine levels. Therefore, as previously mentioned in remarks 5.8 and 5.14, the overall white noise sampling complexity is asymptotically linear with respect to the number of cells of the FEM mesh considered even in the case in which we are not using compactly supported Haar wavelets (cf. Remark 5.2). In the other cases of the MLMC theorem (Theorem 2.1), it might be detrimental to cap the Haar level and compactly supported Haar wavelets are to be preferred in case we cannot afford the additional logarithmic term in the complexity estimate.

## 5.7 Chapter summary and conclusions

In this chapter we presented a novel algorithm to efficiently compute the action of white noise and sample Matérn fields within a QMC and MLQMC framework. This algorithm retains the computational efficiency of the MLMC case (cf. Chapter 4) and still enforces the required multilevel coupling in a non-nested mesh hierarchy. The numerical results show that our technique works well in practice, that the convergence orders observed agree with the theory and that MLQMC outperforms MLMC and has a better cost complexity in the pre-asymptotic regime. As in the MLMC case, we remark that the sampling technique presented extends naturally to any application in which spatial white noise realisations are needed within a finite element framework provided that the solution is smooth enough. An open problem is the derivation of a close-form expression for the optimal number of samples on each MLQMC level and for the optimal Haar level hierarchy, but we leave this to future research.

# Chapter 6

## Uncertainty quantification of tracer distribution in the brain using random interstitial fluid velocity fields

The contents of this chapter are novel unless otherwise indicated. Part of this chapter has been adapted from the paper by C., Vinje and Rognes [44], which is to appear in Fluids and Barriers of the CNS and was written in collaboration with Vegard Vinje, a PhD student/employee from the author’s InFoMM partner company, Simula Research Laboratory. Within this paper, the biological and medical aspects were Vinje’s responsibility, while the numerical and stochastic aspects were the responsibility of the thesis’ author who produced most of the numerical results including all the Monte Carlo simulations. The mathematical modelling and the write-up was done jointly. The contents of this chapter that do not appear in the paper, including the extension to multilevel and quasi Monte Carlo, are novel contributions and were produced by this thesis’ author. These research extensions were supported by the NOTUR grant NN9316K.

In this chapter we have a two-fold objective: we aim to gain insights of clinical relevance and we want to test and compare the methods presented in chapters 4 and 5. While working on the solution of the problem of interest with the co-authors of [44], it became apparent that the initial results obtained with standard Monte Carlo were already accurate enough to make conclusions of clinical relevance (only a few digits of accuracy are needed in practice). Given the fact that more advanced Monte Carlo methods would have still obtained the same conclusions, we decided that, for the sole objective of comparing these methods, considering, for the sake of simplicity, a variant of the forward UQ problem of interest was sufficient.

### 6.1 Introduction

Over the last decade, there has been a significant renewed interest in the waterscape of the brain; that is, the physiological mechanisms governing cerebrospinal fluid (CSF) and interstitial fluid (ISF) flow in (and around) the brain parenchyma. A

number of new theories have emerged including the glymphatic system [93, 96] and the intramural periarterial drainage (IPAD) theory [5, 37] (cf. Section 1.3.2), along with critical evaluations [14, 90, 159]. A great deal of uncertainty and a number of open questions relating to the roles of diffusion, convection and clearance within the brain parenchyma remain.

The glymphatic hypothesis is still far from being established as some aspects of ISF movement are still not well understood. According to the glymphatic hypothesis a bulk flow (which is a consequence of advection-dominated fluid movement) of ISF has been proposed to occur through the brain parenchyma from paraarterial to the paravenous spaces [96]. This directional fluid movement has been named the glymphatic system, with bulk flow being a mechanism for effective waste clearance from the brain parenchyma, which might be related to Alzheimer disease (cf. Chapter 1). The glymphatic hypothesis is strongly dependent on the assumption that a strong bulk flow is present within the brain.

The exchange between CSF and ISF are believed to occur along small fluid filled spaces surrounding large penetrating arteries in the brain parenchyma known as paravascular spaces (PVS; recall Figure 1.2) [93, 145]. However, while evidence of influx of tracer along paravascular spaces seems evident, the outflux route is more debated. Carare et al. [37] found evidence of solutes draining from the brain parenchyma along basement membranes of capillaries and arteries, going in the opposite direction of blood flow and possible PVS fluid movement. This flow is however not facilitated by arterial pulsations [48], but by the movement of smooth muscle cells.

One of the main experimental tools used to study brain fluid flow is magnetic resonance imaging (MRI), which is used to track the movement of injected tracer particles. In fact, all of the in-vivo studies that we mentioned so far have used contrast agents to track the movement of fluid within the intracranial space. However, this might not be the most accurate approach: injection of tracer at rates as low as 1  $\mu\text{L}/\text{min}$  can already cause a significant increase of local intracranial pressure [172], which may lead to pressure gradients driving bulk flow. Furthermore, net movement of tracer does not necessarily require bulk flow and can be explained with dispersion, a combination of diffusion and local mixing [9]. It is therefore not clear whether the recorded evidence for the existence of a glymphatic pathway could have been affected by measurement errors.

For this reason, other experimental methodologies have been used [85]. An alternative avenue of investigation is offered by computational models. Numerical simulations have been employed to assess the possibility of bulk flow within the parenchyma [92], finding it to be dominated by diffusion and not by bulk flow (advection). Smith et al. [159] came to the same conclusion by experimenting with very low infusion rates.

Even though computational models can distinguish between diffusion and bulk flow or advection, there are a vast number of unknown parameters and boundary conditions needed to accurately predict the movement of ISF in the brain parenchyma. For instance, the permeability of brain tissue used in computational models varies from  $10^{-10}$  to  $10^{-17} \text{ m}^2$  [79, 92]. Because the permeability is directly linked to the Darcy fluid velocity in these models, this parameter choice could result in a difference

of 7 orders of magnitude in predicted ISF flow velocities. In addition, CSF dynamics vary between subjects [16] and human CSF production has been reported to increase in the sleeping state [128] which may alter ISF flow.

In a recent review, Abbott and colleagues [2] concluded that bulk flow is likely to be restricted to PVS and possibly white matter tracts, but not in the gray matter. Earlier studies have reported a bulk flow velocity of less than  $1 \mu\text{m/sec}$ , while Bedussi and colleagues [17] recently reported average net bulk flow of  $17 \mu\text{m/sec}$ , the latter study only focusing on PVS. Nevertheless, since tracer movement in *in vivo* studies does not necessarily directly reflect fluid flow, and because model parameters are uncertain with large variability, the exact velocity field governing ISF flow in the brain remains unknown.

For this reason it has recently been pointed out that there is an overarching need to reduce uncertainty when characterizing the anatomy and fluid dynamics parameters in models considering the glymphatic circulation [155]. This study is a first step in this direction: we aim to rigorously quantify how the aforementioned uncertainties in the physiological parameters and in ISF flow affect the spread of a tracer from the subarachnoid space (SAS) into the brain parenchyma. We assume movement of tracer in the brain parenchyma to occur by diffusion and/or convection. To account for uncertainty and variability, we circumvent the lack of precise parameter values by modelling velocity and diffusivity as Matérn stochastic fields. We then set up a convection-reaction-diffusion model with these stochastic fields as coefficients and quantify the uncertainty in the model prediction via a combination of the Monte Carlo and Gaussian field sampling methods presented in the previous chapters. We employ standard MC, MLMC and QMC, but we leave the comparison with MLQMC to future work.

More specifically, we model the contrast MRI study performed by Ringstad et al. [146] assessing glymphatic function in the human brain. We derive a baseline convection-reaction-diffusion PDE, which we then extend by modelling its coefficients according to different hypotheses on CSF flow and clearance, including the relative importance of diffusion, the glymphatic system, possible capillary absorption, and uncertainty within each hypothesis. A total of five different coefficient models (from now on just “models”) were investigated, each with random coefficients. For each model, we compute the expected values, the probability density functions, and 99.73% prediction intervals for different functionals of interest of the tracer concentration. The results reported in the study by Ringstad et al. are compared with the range of uncertainty in our model, solved on a real-life 3D brain FEM mesh derived via MRI. We find that although the uncertainty associated with diffusion yields great variability in tracer distribution, diffusion alone is not sufficient to explain transport of tracer deep into the white matter as seen in experimental data. A glymphatic velocity field may increase tracer concentration deep into the brain, but only when adding a directional structure to the glymphatic circulation.

Replacing PDE parameters subject to uncertainty with spatially correlated random fields is a common modelling choice in the UQ literature [39, 41, 166] and Monte Carlo methods have been successfully used in biology to quantify how uncertainty in model input propagates to uncertainty in model output. However, these methods

have mainly been applied to simulations of the cardiovascular system [24, 142] and, to our knowledge, there has only been one study in which Monte Carlo methods have been used for UQ in brain modelling [86]. To the author’s knowledge, there has been no previous work in which multilevel Monte Carlo or QMC methods have been applied to brain simulation and this is the first study in which forward UQ is applied to simulations of tracer transport within the brain parenchyma.

The outline of the chapter is as follows. In Section 6.2 we derive 5 stochastic models for brain tracer movement inspired by the MRI study of Ringstad et al. [146]. We describe the numerical strategy employed for their solution in section 6.3. In Section 6.4 we present the simulation results and their clinical implications. We conclude the chapter with Section 6.5 where we compare the performance of our MLMC and QMC sampling strategies when applied to two variants of the models of interest.

## 6.2 Stochastic modelling of brain tracer concentration

We model the MRI-study of Ringstad et al. [146]. In their experiments, 0.5 mL of 1.0 mmol/mL of the radioactive tracer gadobutrol was injected in the spinal canal (intrathecally) of 15 hydrocephalus<sup>1</sup> patients and eight reference subjects. In this study we only consider the healthy patient case. The localisation of the tracer was found with MRI at 4 different time periods, at 1, 3, 4.5, and 24 hours following the injection. After 3 hours, tracer was localised in the lower region of the cranial SAS, and had started to penetrate into the brain parenchyma of the reference subjects. The following day it had spread throughout the brain tissue. Tracer was found to penetrate along large arteries in all study subjects, and a low proportion of tracer was found in the upper regions of the brain.

### 6.2.1 The baseline model

We consider the following partial differential equation with random coefficients to model transport of tracer in the brain parenchyma under uncertainty: find the tracer concentration  $c = c(t, \mathbf{x}, \omega)$  for  $\mathbf{x} \in G$ ,  $\omega \in \Omega$  and  $t \geq 0$  such that

$$\dot{c}(t, \mathbf{x}, \omega) + \nabla \cdot (\mathbf{v}(\mathbf{x}, \omega)c(t, \mathbf{x}, \omega)) - \nabla \cdot (D^*(\mathbf{x}, \omega)\nabla c(t, \mathbf{x}, \omega)) + rc(t, \mathbf{x}, \omega) = 0. \quad (6.1)$$

Here,  $G \subset \mathbb{R}^3$  is the brain parenchyma domain, the superimposed dot represents the time derivative,  $D^*$  is the effective diffusion coefficient of the tracer in the tissue (depending on the tracer free diffusion coefficient and the tissue tortuosity) [127],  $\mathbf{v}$  represents a convective fluid velocity and  $r \geq 0$  is a drainage coefficient potentially representing e.g. capillary absorption [130] or direct outflow to lymph nodes [146]. We assume that the parenchymal domain contains no tracer initially:  $c(0, \mathbf{x}, \omega) = 0$ .

---

<sup>1</sup>A brain disease in which CSF accumulates within the brain causing migraine, incontinence, behavioural changes and mild dementia.

Model	$D^*$	$v$	$r$
D1	Random variable	0	0
D2	Random field	0	0
V1	Constant	Random influx and outflux field	0
V2	Constant	Model V1 + directional velocity field	0
V3	Constant	Random influx field	$r > 0$
RF1	Random field	Model V1 + directional velocity field	0
RF2	Random field	Random influx field	$r > 0$

Table 6.1: *Summary of stochastic model variations with effective diffusion coefficient  $D^*$ , convective fluid velocity  $v$ , and drainage coefficient  $r$  in (6.1). Models RF1 and RF2 are only used for comparing the Monte Carlo methods presented in the previous chapters and have less clinical relevance.*

To investigate and compare different hypotheses for parenchymal ISF flow and tracer transport, we consider 5 stochastic model variations of (6.1) including two models with random diffusion (model D1 and D2) and three models with stochastic velocity fields (models V1, V2, and V3). The diffusion-only models D1 and D2 correspond to negligible ISF bulk flow in the parenchyma and the absence of capillary absorption or other direct outflow pathways. For the velocity models (V1, V2 and V3), we consider a constant diffusion coefficient in order to isolate the effects of the stochastic velocity fields. We also consider two more complicated models with random diffusivity and velocity, models RF1 and RF2, that we use to test the different sampling techniques presented in the previous chapters. A summary of the models is presented in Table 6.1, while the mathematical modelling aspects are described in further detail in the following sections.

## Domain and geometry

We define the computational domain  $G$  as the union of white and gray matter (cf. Section 1.3) from the generic Colin27 human adult brain atlas FEM mesh [58] version 2 (Figure 6.1a). This domain includes the cerebellum. The levels of the brain regions from which measurements were taken in the study by Ringstad et al. (foramen magnum, the sylvian fissure and the precentral sulcus) are well represented by the  $x_3$ - or  $z$ -coordinates -0.1, 0 and 0.1 m, respectively. The plane  $z = 0$  corresponds approximately to the level of the lateral ventricles (cf. Figure 6.1).

## Boundary conditions modelling tracer movement in the SAS

Let  $\partial G$  be the boundary of  $G$  and let  $\partial G = \partial G_S \cup \partial G_V$ , with  $\partial G_S$  representing the interface between the brain parenchyma and the subarachnoid space (SAS), and  $\partial G_V$  representing the interface between the brain parenchyma and cerebral ventricles,

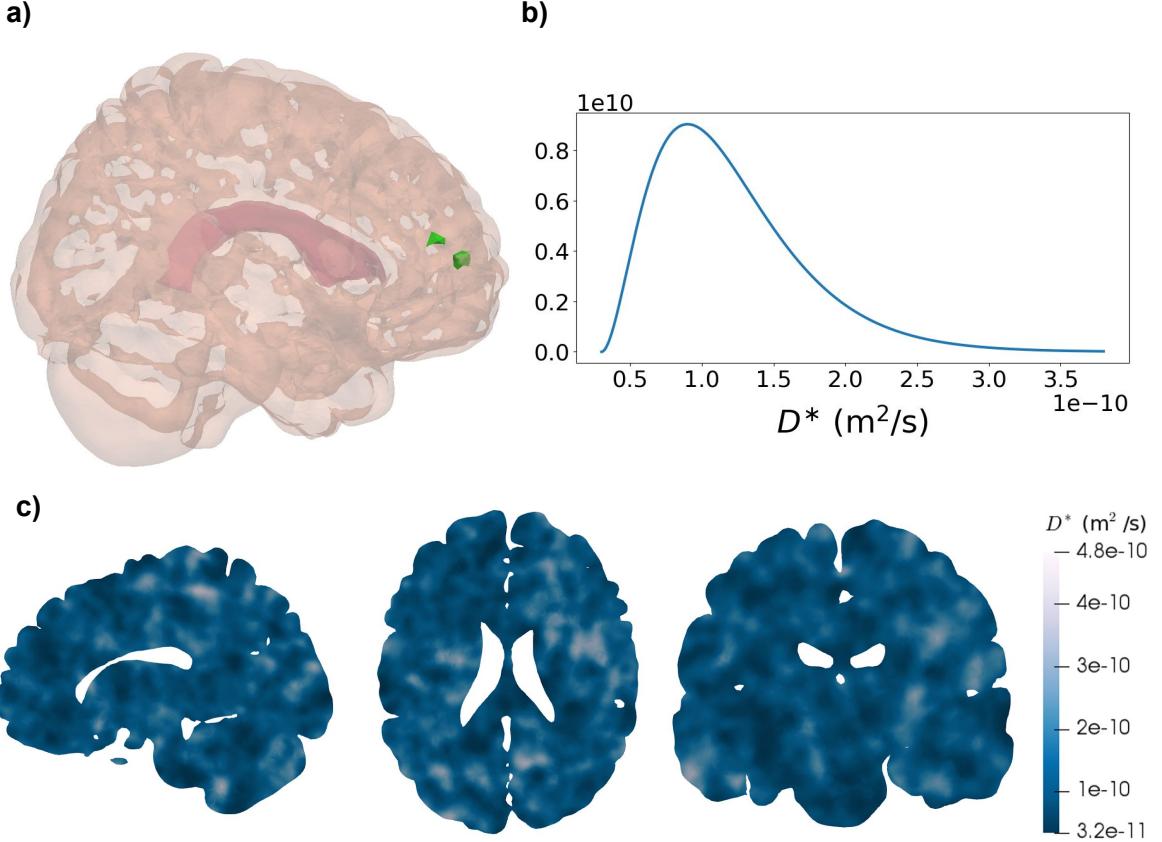


Figure 6.1: Computational domain and random diffusion coefficient. a) The computational domain representing the brain parenchyma including the cerebellum. The interior lateral ventricles are marked (dark pink) in the central region of the domain. Two smaller regions of interest  $S_g$  and  $S_w$ , in the gray and white matter respectively, are marked in green (leftmost region:  $S_w$ , rightmost region:  $S_g$ ). b) Assumed probability distribution of the homogeneous effective diffusion coefficient  $D^*$  used in model  $D1$ . The expected value  $E[D^*]$  is  $1.2 \times 10^{-10} \text{ m}^2/\text{s}$ . c) Sample of the heterogeneous effective diffusion coefficient (sagittal, axial and coronal slices ordered from left to right) modelled as a random field and used in model  $D2$ .

respectively. We consider the following boundary conditions for (6.1):

$$c = g(t, \mathbf{x}, \omega) \text{ on } \partial G_S, \quad (6.2)$$

$$D^* \nabla c \cdot \mathbf{n} = 0 \text{ on } \partial G_V, \quad (6.3)$$

where  $\mathbf{n}$  is the unit normal vector pointing outward from  $\partial G$ . In particular, we assume that a tracer concentration is given at the SAS interface (6.2) and no ventricular outflux (6.3).

The boundary condition (6.2) models the movement of tracer starting from the lower cranial SAS (recall that the tracer was injected in the spinal canal) and traveling upward in the CSF surrounding the brain as observed in the study by Ringstad et

al. [146]. In particular, we let

$$\begin{aligned} g(t, \mathbf{x}, \omega) &= c_{\text{CSF}}(t, \omega) h(t, \mathbf{x}), \\ h(t, \mathbf{x}) &= \left( 0.5 + \frac{1}{\pi} \arctan(-a(x_3 - z_0 - v_z t)) \right), \end{aligned} \quad (6.4)$$

for  $\mathbf{x} = (x_1, x_2, x_3)$ . Here, at time  $t$ ,  $c_{\text{CSF}}(t)$  is the average tracer concentration in the SAS and depends on  $c$  (as detailed below), while  $h(t, \mathbf{x})$  represents its spatial distribution.

The expression for  $h$  is based on the following considerations. We assume that the diffusive and/or convective movement of tracer from the spinal to the cranial SAS over time is known, and we thus model  $h(t, \mathbf{x})$  as a smooth step function upwards (in the  $x_3$ - or  $z$ -direction). In (6.4),  $v_z$  represents the speed of tracer movement upwards in the SAS, and  $a$  reflects the rate of change of tracer concentration from the lower to the upper cranial SAS. Finally, we assume that at time  $t = 0$ , the tracer has spread up to a relative distance of  $z_0$  from the lateral ventricles. This specific expression for  $h(t, \mathbf{x})$  and the values of parameters  $a$ ,  $z_0$  and  $v_z$  are based on the spread of tracer seen in the MR-images in the study by Ringstad et al. [146]. In particular, we use  $a = 20 \text{ m}^{-1}$ ,  $v_z = 1.5 \times 10^{-5} \text{ m/sec}$  and  $z_0 = -0.2 \text{ m}$ . These parameters were chosen to match time to peak in three different regions in the CSF space in reference individuals [146].

To derive an expression for  $c_{\text{CSF}}$  in (6.4), we consider the conservation of tracer mass. We model the spread of  $n_0 = 0.5 \text{ mmol}$  tracer in the CSF, assuming a volume of  $V_{\text{CSF}} = 140 \text{ mL}$  CSF in the human SAS and ventricles [177]. The average concentration in the SAS right after injection is thus  $c_{\text{CSF}}(0) = 0.5 \text{ mmol}/140 \text{ mL} = 3.57 \text{ mol/m}^3$ . At any given time, we assume that the total amount of tracer in the brain and in the SAS plus or minus the tracer absorbed or produced stays constant in time, and is equal to the initial amount  $n_0 = 0.5 \text{ mmol}$  (almost surely). Applying a conservation of tracer molecules principle gives

$$\int_G c(t, \mathbf{x}, \omega) \, d\mathbf{x} + c_{\text{CSF}}(t, \omega) V_{\text{CSF}} + \int_0^t \int_G r c(\tau, \mathbf{x}, \omega) \, d\mathbf{x} \, d\tau = n_0. \quad (6.5)$$

Solving for  $c_{\text{CSF}}$ , we thus obtain

$$c_{\text{CSF}}(t, \omega) = \frac{1}{V_{\text{CSF}}} \left( n_0 - \int_G c(t, \mathbf{x}, \omega) \, d\mathbf{x} - \int_0^t \int_G r c(\tau, \mathbf{x}, \omega) \, d\mathbf{x} \, d\tau \right). \quad (6.6)$$

### Quantities of interest

To evaluate the speed and characteristics of tracer movement into and in the brain parenchyma, we consider a set of functionals describing different output quantities of interest. To quantify the overall spread of tracer in the gray and white matter, we consider the total amount of tracer in the gray matter  $Q_g$  and in the white matter  $Q_w$  at time points  $\tau$ :

$$Q_g(\omega) = \int_{D_g} c(\tau, \mathbf{x}, \omega) \, d\mathbf{x}, \quad Q_w(\omega) = \int_{D_w} c(\tau, \mathbf{x}, \omega) \, d\mathbf{x}. \quad (6.7)$$

We pay particular attention to the times  $\tau \in \{3, 5, 8, 24\}$  as they correspond to the times considered in the study by Ringstad et al. To further differentiate, we also defined two localised functionals at each time  $\tau$ : the average tracer concentration  $q_g$  in a small subregion of the gray matter  $S_g$  and analogously  $q_w$  for a small subregion of the white matter  $S_w$ :

$$q_g = \frac{1}{V_g} \int_{S_g} c(\tau, \mathbf{x}, \omega) \, d\mathbf{x}, \quad q_w = \frac{1}{V_w} \int_{S_w} c(\tau, \mathbf{x}, \omega) \, d\mathbf{x}, \quad (6.8)$$

where  $V_g$  and  $V_w$  is the volume of the gray and white matter subregions, respectively. The size and relative location of the subregions  $S_g$  and  $S_w$  within the computational domain are illustrated in Figure 6.1a. To further quantify the speed of propagation, we define the white matter activation time  $F_w$ :

$$F_w(\omega) = \{\min t \mid \int_{\Omega_w} c(t, \mathbf{x}, \omega) \, d\mathbf{x} / n_0 > \Upsilon_1\}, \quad (6.9)$$

where  $n_0$  is the total amount of tracer injected into the SAS (0.5 mmol) and  $\Upsilon_1$  is a given percentage. Given the time course of the expected tracer distribution to the white matter [147], we here chose  $\Upsilon_1 = 10\%$ . Finally, we also define the analogous regional (white matter) activation time

$$f_w(\omega) = \{\min t \mid \frac{1}{V_w} \int_{S_w} c(t, \mathbf{x}, \omega) \, d\mathbf{x} > \Upsilon_2\}, \quad (6.10)$$

where we choose  $\Upsilon_2 = 10^{-3} \text{ mol/m}^3$  (choice inspired by [147]).

For plotting the boundary tracer concentration over time, we define three axial planes along the z-axis ( $z = -0.1, 0, 0.1 \text{ m}$ ) to represent the level of three different regions of interests within the brain (foramen magnum, sylvian fissure and precentral sulcus, respectively).

### 6.2.2 Stochastic diffusion modelling

The parenchymal effective diffusion coefficient of a solute, such as e.g. gadobutrol, is heterogeneous [170] (varies in space) and individual-specific (varies from individual to individual). To investigate the effect of uncertainty in the diffusion coefficient, we consider two approaches: first, to model the diffusion coefficient as a random variable and second, to model the diffusion coefficient as a random field, thus allowing for tissue heterogeneity. Both approaches are described in further detail below.

#### Effective diffusion coefficient modelled as a random variable

First, we consider the simplifying but common assumption that the effective diffusion coefficient is spatially homogeneous:  $D^*(\omega) \in \mathbb{R}_+$  a.s. We account for the uncertainty in its value by modelling it as a random variable:

$$D^*(\omega) = 0.25 \times D_{\text{Gad}}^* + D_\gamma^*(\omega), \quad (6.11)$$

where  $D_{\text{Gad}}^* = 1.2 \times 10^{-10} \text{ m/s}^2$  is a fixed parenchymal gadobutrol diffusivity [147] and  $D_\gamma^*(\omega)$  has a Gamma distribution with shape  $k = 3$  and scale  $\theta = 0.75 \times D_{\text{Gad}}^*/k$ . The choice of shape and scaling parameters ensures that (i) the diffusion coefficient is positive, (ii) its expected value matches reported values of parenchymal gadobutrol diffusivity [147], and (iii) its variability allows for values up to 2–3 times larger or smaller than the average with low probability. The last modelling choice reflects diffusivity values in the range  $1\text{--}10 \times 10^{-10} \text{ m/s}^2$  in agreement with previous reports [127]. The probability distribution of  $D^*$  is shown in Figure 6.1b-c.

### Effective diffusion coefficient modelled as a random field

In order to represent spatial heterogeneity in the diffusion coefficient, we next model  $D^*$  as a continuous random field. Again, we set

$$D^*(\mathbf{x}, \omega) = 0.25 \times D_{\text{Gad}}^* + D_f^*(\mathbf{x}, \omega), \quad (6.12)$$

where  $D_f^*$  now is a random field such that for each fixed  $\mathbf{x} \in G$ ,  $D_f^*(\mathbf{x}, \omega)$  is a gamma-distributed random variable with the same parameters as  $D^*(\omega)$  in (6.11). To enforce continuity and to easily sample the random field from its distribution, we draw samples of  $D_\gamma^*$  by first sampling a Matérn field  $X(\mathbf{x}, \omega)$  using the techniques presented in chapters 4 and 5, and then transform it into a gamma random field by using a copula [125]. This consists in setting  $D_\gamma^*(\mathbf{x}, \omega) = F^{-1}(\Phi(X(\mathbf{x}, \omega)))$ , where  $F^{-1}$  is the inverse CDF of the target (gamma) distribution,  $\Phi$  is the CDF of the standard normal distribution and  $X(\mathbf{x}, \omega)$  is a standard (zero mean, unit variance) Matérn field with smoothness parameter  $\nu = 2.5$  and correlation length  $\lambda = 0.01 \text{ m}$ , cf. (2.27). Note that  $\Phi$  maps any standard normal random variable to a standard uniform random variable and that  $F^{-1}$  maps any standard uniform random variable to the target distribution, hence the function  $F^{-1}(\Phi(x))$  maps standard random variables to the target gamma distribution. Samples of  $D_\gamma^*(\mathbf{x}, \omega)$  obtained this way will preserve the same Spearman correlation and smoothness properties of  $X(\mathbf{x}, \omega)$ , but will present a different covariance structure [125] ( $F^{-1}(\Phi(x))$  is non-linear, but smooth and monotone).

### 6.2.3 Stochastic velocity modelling

In what follows we introduce three different models for the velocity field, each representing a different hypothesis regarding intraparenchymal ISF/CSF movement. We emphasise that each model represent a homogenised velocity field averaged over physiological structures.

#### Glymphatic velocity model: arterial influx and venous outflux

To define a stochastic homogenised velocity model representing the glymphatic pathway, we assume that ISF follows separate inflow and outflow routes: entering the brain along paraarterial spaces and exiting along paravenous spaces [96]. We further suggest that

1. Substantial changes within the velocity field happen after a distance proportional to the mean distance between arterioles and venules.
2. The blood vessel structure is random and independent from the position within the parenchyma in the sense that the presence of paraarterial or paravenous spaces are equally likely at any point in space. Mathematically, this assumption requires the expected value of each of the velocity components to be zero.
3. The velocity field varies continuously in space and is divergence-free ( $\nabla \cdot \mathbf{v} = 0$ ), i.e. no CSF/ISF leaves the system e.g. through the bloodstream.
4. We set the expected velocity magnitude  $|\mathbf{v}|_2 = \sqrt{v_x^2 + v_y^2 + v_z^2}$  to be  $v_{\text{avg}} = 0.17 \mu\text{m/s}$  and we allow for up to 2-3 times larger and up to 10 times smaller values with low probability [127].

Although ISF/CSF velocities in paravascular regions may be higher [122] than what we propose, the velocity field here models an averaged bulk flow over a larger area (comprised of e.g. PVS and adjacent tissue). Bulk flow velocities in rats have been reported to be in the range of approximately 0.1-0.24  $\mu\text{m/s}$  [1, 127].

To address these stipulations, we define the stochastic glymphatic circulation velocity field

$$\mathbf{v}(\mathbf{x}, \omega) = v_{\text{avg}} \cdot \bar{\eta}(\lambda) \sqrt{U(\omega)} \left( \nabla \times \begin{bmatrix} X(\mathbf{x}, \omega) \\ Y(\mathbf{x}, \omega) \\ Z(\mathbf{x}, \omega) \end{bmatrix} \right), \quad (6.13)$$

where  $\bar{\eta}(\lambda) = \lambda/\sqrt{8}$  is a scaling constant chosen such that the magnitude of  $\mathbf{v}$  satisfies  $\mathbb{E}[|\mathbf{v}|_2^2]^{1/2} = v_{\text{avg}}$  (we omit the mathematical derivation of this constant),  $U(\omega)$  is a standard uniform random variable and  $X(\mathbf{x}, \omega)$ ,  $Y(\mathbf{x}, \omega)$  and  $Z(\mathbf{x}, \omega)$  are standard i.i.d. Matérn fields with  $\nu = 2.5$  and correlation length  $\lambda = 1020 \mu\text{m}$ . A sample of the glymphatic circulation velocity field together with the velocity magnitude distribution is shown in Figure 6.2a-b.

The factor  $\sqrt{U(\omega)}$  is an ad-hoc term to enforce the variability requirement defined by point 4) above. The use of Matérn fields enforces spatial variability in a continuous manner and taking the curl operator ensures that the resulting velocity is divergence-free, hence addressing point 3). It can be proven that the field within the brackets in (6.13) is still Gaussian, has zero mean (hence satisfies point 2) and has the same correlation length as the original Matérn fields, albeit it presents a slightly different covariance structure. To see this, note that the partial derivative  $\partial X/\partial x_i$  of a zero-mean Gaussian field  $X(\mathbf{x}, \omega)$  with a twice differentiable covariance  $C(\mathbf{x}, \mathbf{y})$  is still a zero-mean Gaussian field with covariance given by  $\partial^2 C(\mathbf{x}, \mathbf{y})/(\partial x_i \partial y_i)$  (cf. section 2.3 in [3]). The curl components are therefore just sums of independent Gaussian fields, and hence Gaussian as well. Their covariance function is proportional to a second partial derivative of the Matérn covariance function (2.27) and hence the correlation length is preserved, although the resulting covariance is not Matérn anymore and the components of  $\mathbf{v}$  cease to be independent.

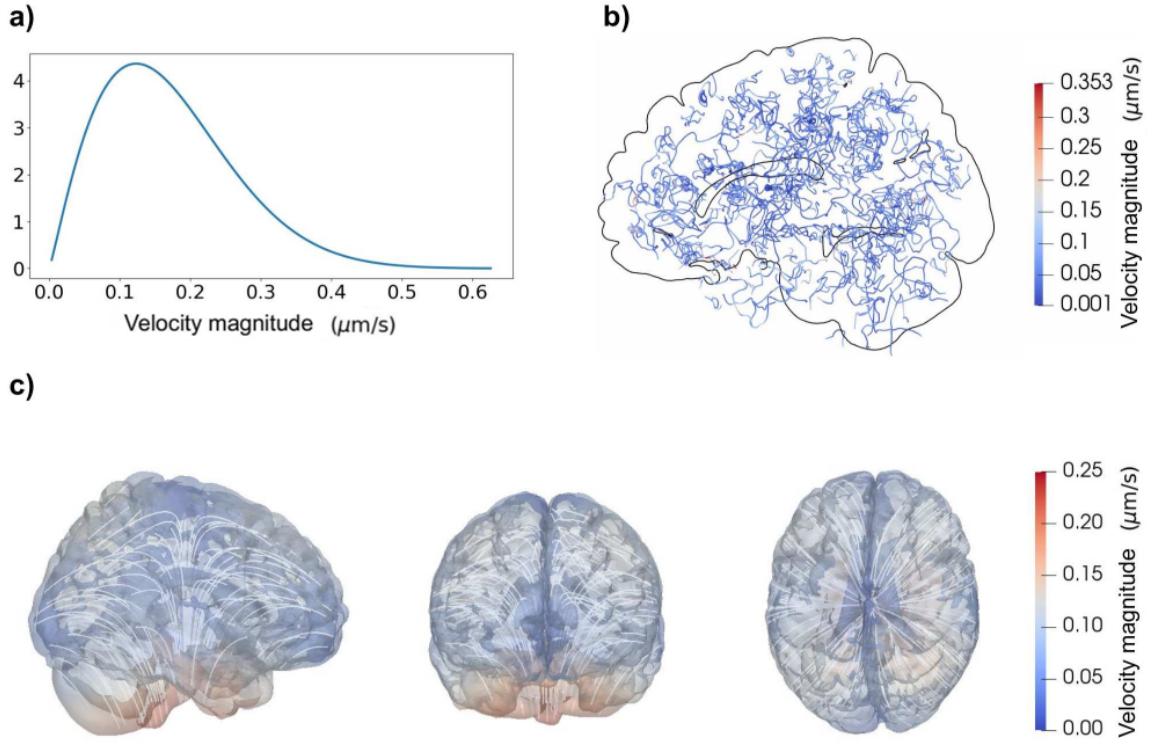


Figure 6.2: *Stochastic aspects of the glymphatic circulation velocity fields (models V1 and V2).* a) Probability density of the glymphatic circulation velocity magnitude  $|\mathbf{v}|_2$  cf. (6.13). b) Streamlines of a corresponding velocity field sample. c) Velocity magnitude and streamlines for the directional velocity field  $v_{\text{dir}}$  as given by (6.14). The flow field is assumed to follow cardiovascular pulses upwards along the brain. After entering the deeper parts of the brain, the bulk flow spreads out at reduced velocity. From left to right: sagittal, coronal and transverse view.

The choice of correlation length was guided by the following considerations. The mean distance between arterioles and venules was reported to be  $280 \mu\text{m}$  in rhesus monkeys [4], although the value  $250 \mu\text{m}$  has been used as a representative distance in humans in recent modeling papers [97, 144]. We estimated the mean distance in humans by considering differences in brain and artery size between monkey and human (Table 6.2). We find a factor close to 2 between CCA and arteriole diameter, while a similar ratio was found for the cube root of the brain mass. Thus, the correlation length should be greater than  $250 - 560 \mu\text{m}$  to address point 1) above. Combining these physiological considerations with the corresponding requirements on the numerical resolution, we let  $\lambda = 1020 \mu\text{m}$ .

Species	Brain mass (g)	$d_{\text{CCA}}$ (mm)	$d_A$ ( $\mu\text{m}$ )	$\Delta_{\text{AV}}$ ( $\mu\text{m}$ )
Mouse	0.3 [148]	0.47 [111]	25 [94]	40*
Monkey	88 [148]	3.5[180]	35.5 [4]	280 [4]
Human	1350 [148]	6.3 [106]	40-250 [11]	1020*

Table 6.2: *Brain-related parameters of three species.* \*: Estimated values.  $d_{\text{CCA}}$ : diameter of the common carotid artery,  $d_A$ : arteriole diameter  $\Delta_{\text{AV}}$ : distance between arteriole and venule.

### Glymphatic velocity model with additional directional velocity field

Above we assumed that the blood vessel distribution was independent of the spatial position within the parenchyma and that bulk flow from arterial to venous PVS occurs on a small length scale proportional to the mean distance between arterioles and venules. However, transport of tracer might also happen on a larger length scale along larger vascular structures present in given physical regions (such as e.g. circle of Willis<sup>2</sup>). As CSF is hypothesised to enter the brain along penetrating arteries, the direction of cardiac pulse propagation may induce a directionality of the (glymphatic) ISF circulation as well. The cardiac pulse follows the vessels of larger arteries entering the brain from below, and from there spreads out almost uniformly [104, 143]. The pulses also seem to traverse deep gray matter regions on the way up towards the ventricles.

To model such behavior, we introduce a directional velocity field  $v_{\text{dir}}$ , with characteristics qualitatively similar to what is described in the literature [104, 143]:

$$\mathbf{v}_{\text{dir}}(\mathbf{x}) = -v_f \begin{pmatrix} \arctan(15x_1)(|x_1| - 0.1) \\ \arctan(15x_2)(|x_2| - 0.1) \\ -0.9x_3 + 0.06 - \sqrt{x_1^2 + x_2^2} \end{pmatrix}, \quad (6.14)$$

where  $v_f = 2 \times 10^{-6}$  m/s. For a plot of  $\mathbf{v}_{\text{dir}}$ , see Figure 6.2c. The velocity field  $\mathbf{v}_{\text{dir}}$  induces a net flow out of the parenchyma at the very low rate of 0.007 mL/min. We superimpose this deterministic directional velocity field on the stochastic glymphatic circulation velocity field to define the stochastic glymphatic directional velocity field:

$$\mathbf{v}(\mathbf{x}, \omega) = \mathbf{v}_{\text{V1}}(\mathbf{x}, \omega) + \mathbf{v}_{\text{dir}}(\mathbf{x}), \quad (6.15)$$

where  $\mathbf{v}_{\text{V1}}$  is given by (6.13). This velocity model thus takes into account both the “randomness” of small arteries, but also the “deterministic” presence of large arteries and possibly other structures of blood flow propagation [104, 143].

---

<sup>2</sup>A large interconnection of vascular channels that provides oxygenated blood to the brain.

### Capillary filtration model V3: arterial inflow with a homogeneous sink throughout the brain

Several independent studies demonstrate that CSF may enter the brain parenchyma along spaces surrounding penetrating arteries [5, 18, 96, 122]. However, the glymphatic outflux hypothesis of an advection-dominated flow of CSF through the extracellular space and recirculation into the SAS through paravenous spaces has been severely questioned [5, 90, 92, 158]. As a variation, we here therefore also consider a stochastic velocity model representing paraarterial influx without a direct return route to the CSF. Instead, we assume that ISF/CSF is drained inside the brain parenchyma along some alternative outflux pathway. This pathway may include the capillaries or separate spaces along the PVS directly into the lymphatic system at cervical lymph nodes.

In light of this, we consider the following alternative velocity assumptions. (1) There is a net flow of CSF into the brain and (2) ISF is cleared within the parenchyma via some, here unspecified, route. For instance, it has been proposed that production and absorption is present all over the CSF system and that capillaries and ISF continuously exchanges water molecules [131]. However, drainage of large molecules through this route is unlikely [90]. It has also been reported that lymph vessels may be capable of also draining larger molecules from brain tissue into deep cervical lymph nodes, possibly through paravenous spaces [10]. In addition, other outflow routes may exist [165].

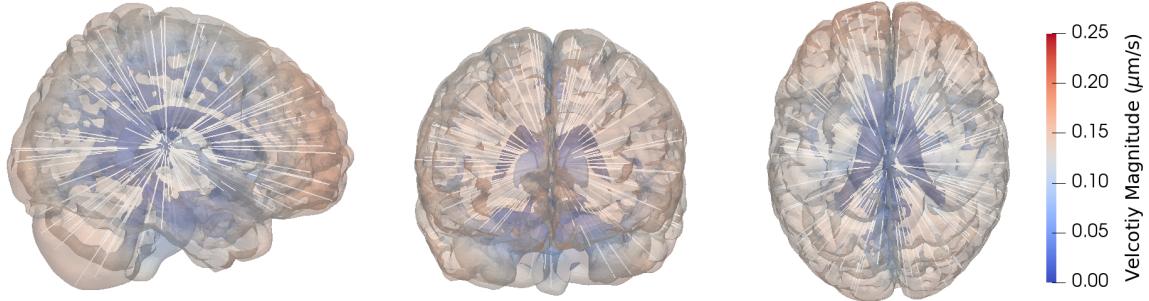


Figure 6.3: *Sample model V3 velocity field. Velocity magnitude and streamlines for the velocity field as given by (6.16). Flow is assumed to occur from the cortex towards the ventricles with reduced velocity along the way due to clearance. From left to right: sagittal, coronal and transverse view.*

To address these assumptions, we define a stochastic arterial inflow velocity field as a radially symmetric field pointing inwards from the SAS interface to the brain region around the lateral ventricle. This central region is modelled in what follows as a sphere of radius  $R = 8$  cm and center given by  $\mathbf{x}_c$  in the lateral ventricles.

Mathematical experimentation lead to the following *ansatz* for such velocity:

$$\mathbf{v}(\mathbf{x}, \omega) = \bar{v}(\omega) \exp\left(-\frac{3(R - \|\mathbf{x} - \mathbf{x}_c\|)^2}{R^2 - (R - \|\mathbf{x} - \mathbf{x}_c\|)^2}\right) (\mathbf{x}_c - \mathbf{x}), \quad (6.16)$$

where  $\bar{v}(\omega)$  is a gamma random variable chosen such that the probability distribution of the velocity magnitude is comparable to that of the glymphatic circulation velocity defined by (6.13). The shape parameter  $k = 2$  and the scale parameter is set such that again  $\mathbb{E}[|\mathbf{v}|_2^2]^{1/2} = v_{\text{avg}}$ . Note that in this case, the expected value of the velocity components are non-zero. To satisfy (2), we model the drainage of tracer by setting  $r = 1 \times 10^{-5} \text{ s}^{-1}$ , which typically results in 40% drainage of the injected tracer over 48 hours.

## 6.3 Numerical solution of the stochastic models

### Two different objectives

In this chapter we aim to: 1) solve problem (6.1) to sufficient accuracy, including the estimation of the PDFs of all the output functionals considered, and 2) test and compare the methods presented in chapters 4 and 5. To address point 1, high accuracy is not needed since only a few digits are sufficient to obtain visual convergence of the numerical simulations and draw conclusions of clinical relevance [44]. As a consequence, it became apparent that the initial exploratory results that we obtained with standard Monte Carlo were already accurate enough and that no extra insights would have been found with the use of more advanced methods. We therefore decided that for the sole purpose of making the comparison in point 2, considering, for the sake of simplicity, a variant of the numerical problem at hand was sufficient.

In what follows, we detail the two numerical approaches adopted for the solution of (6.1).

### Weak form and discretisation

Let  $H_S^1(G) := \{s \in H^1(G) \text{ s.t. } s = 0 \text{ on } \partial G_S\}$  and, for  $c, s \in H^1(G)$ , let

$$a(c, s) = (\nabla \cdot (\mathbf{v}c), s) + (D^* \nabla c, \nabla s) + (rc, s). \quad (6.17)$$

After time discretisation, the weak form of (6.1) reads: find  $c^n \in H^1(G)$  such that for all  $s \in H_S^1(G)$  and a.s.,

$$\left( \frac{c^n - c^{n-1}}{\Delta t}, s \right) + \frac{1}{2}(a(c^n, s) + a(c^{n-1}, s)) = 0, \quad (6.18)$$

$$c^0 \equiv 0, \quad c^n = c_{\text{CSF}}^n h(t^n, \mathbf{x}) \text{ on } \partial G_S, \quad (6.19)$$

where  $c^n$  is the FEM approximation to  $c(t^n, \cdot, \omega)$  with  $t^n = n\Delta t$  for  $n = 0, \dots, n_T - 1$  and  $n_T - 1 = T/(\Delta t)$  and  $c_{\text{CSF}}^n$  is an explicit approximation of  $c_{\text{CSF}}(t^n, \omega)$ , defined

in (6.6). We approximate  $c_{\text{CSF}}^n$  explicitly by approximating the time integral in (6.6) with the trapezoidal rule:

$$c_{\text{CSF}}(t^n, \omega) \approx c_{\text{CSF}}^n = \frac{1}{V_{\text{CSF}}} \left( n_0 - \int_G c^{n-1} \, d\mathbf{x} - \frac{\Delta t}{2} \left( 2 \sum_{i=1}^{n-2} \int_G r c^i \, d\mathbf{x} + \int_G r c^{n-1} \, d\mathbf{x} \right) \right). \quad (6.20)$$

Since this boundary condition is discretised explicitly, this results in a first-order scheme in time. A second-order discretisation in time is also feasible, but requires an implicit discretisation of the non-local boundary condition, which can be done by either introducing a Lagrange multiplier or by manually modifying the FEM linear system matrix. The first-order discretisation is accurate enough to address point 1. For point 2, we still keep the main PDE with random coefficients (6.1), but we replace the non-local boundary condition with a Dirichlet boundary condition obtained by solving a deterministic version of equation (6.1) in which each random coefficient is replaced by its expected value. This “overkill” solution is computed by using the first-order in time scheme with an extremely small time step ( $\Delta t = 30 \times 2^{-6}$  min) on the finest mesh available (see later in this section for a description of the meshes and time step sizes used).

*Remark 6.1.* Changing the boundary condition to Dirichlet turns the time discretisation (6.18) into a second order scheme, which in turn allows us to avoid the “bad”  $\gamma > \beta$  case of the MLMC theorem (cf. Theorem 2.1). In fact, a first-order in time and second-order in space discretisation would yield MLMC parameter values of  $\alpha = 2$ ,  $\beta = 4$ ,  $\gamma = 5$ . This is assuming that a perfect multigrid solver and a geometric mesh hierarchy is used so that  $h \propto 2^{-\ell}$  and that we fix the time step so that  $\Delta t \propto 2^{-2\ell}$ . A second-order in space and time discretisation instead gives  $\gamma = \beta = 4$  (for  $h, \Delta t \propto 2^{-\ell}$ ), which is the “borderline” case of the MLMC theorem and has a better total cost complexity estimate. The downside of changing the boundary condition is that this affects the PDF of the solution, which is what we are also trying to estimate as part of point 1. For this reason, we introduce this change only to address point 2 for which the actual probability density of the solution is not important.

We discretise (6.18) in space by using the FEM. Given a FEM approximation subspace  $V_h \subseteq H_S^1(G)$ , the fully discrete weak form of (6.1) reads: find  $c_h^n \in V_h$  such that, for all  $s_h \in V_h$  and a.s.,

$$\left( \frac{c_h^n - c_h^{n-1}}{\Delta t}, s_h \right) + \frac{1}{2} (a(c_h^n, s_h) + a(c_h^{n-1}, s_h)) = 0, \quad (6.21)$$

$$c_h^0 \equiv 0, \quad c_h^n = c_{\text{CSF}}^{n,h} h(t^n, \mathbf{x}) \text{ on } \partial G_S, \quad (6.22)$$

where  $c_{\text{CSF}}^{n,h}$  is given by (6.20) in which  $c^{n-1}$  and  $c^i$  are replaced by  $c_h^{n-1}$  and  $c_h^i$  respectively.

## Meshes and time step sizes used

We discretise the domain  $G$  by using various refinements of the Colin27 human adult brain atlas simplicial mesh [58] (version 2). To address point 1, we use an adaptively refined version of this brain mesh with 1 875 249 vertices and 9 742 384 cells and a time step of  $\Delta t = 15$  min. To address point 2, we construct a multilevel hierarchy in which the coarsest level is given by one uniform refinement of the original brain mesh and the other 2 levels are obtained through uniform refinement. On level  $\ell$ , we fix  $(\Delta t)_\ell = 15 \times 2^{-\ell}$  min. The finest mesh in the hierarchy has 127 672 832 cells and 22 282 705 vertices. Since the Matérn fields need to be sampled on an extended domain, we embed each brain mesh into a mesh of a larger box domain of size sufficiently large to make the domain truncation error negligible (dimensions  $0.16 \times 0.21 \times 0.17$  m<sup>3</sup>) [103]. Each outer box mesh is constructed with the meshing software Gmsh [65] (dev version 4.2.3-git-023542a) so that the corresponding brain mesh is nested within it. Furthermore, the box meshes are graded so that the cell size gradually gets larger away from the brain boundary (Matérn field values are only needed in the brain domain).

## Numerical stabilisation

We choose the FEM subspaces to be spanned by continuous piecewise linear basis functions. We remark that a stabilisation here is not needed since the problem is only mildly convection-dominated, with an upper estimate of the Péclet number of

$$Pe \approx \frac{\hat{L}v_{\text{avg}}}{D_{\text{Gad}}^*} = O(10^2), \quad (6.23)$$

where  $\hat{L} \approx 0.084$  m is half the diameter of the computational domain,  $v_{\text{avg}} = 0.17 \mu\text{m}/\text{s}$ , and  $D_{\text{Gad}}^* = 1.2 \times 10^{-10}$  m/s<sup>2</sup>. Given the small mesh sizes we get a low-probability worst-case cell Péclet numbers of  $\approx 43 \times 2^{-\ell}$  on level  $\ell$  of the MLMC hierarchy used in point 2 and of  $O(10)$  for the mesh used in point 1. In numerical experimentations, numerical instabilities were never observed, not even on the coarsest mesh used. One of the reasons could be that the FEM approximation of the solution of the white noise PDE (2.41) appears to underestimate the second moment of the sampled fields on coarse meshes<sup>3</sup>. This in turn makes the Péclet number smaller since the velocity magnitude directly relates to the second moment of the sampled fields. In fact,  $\mathbb{E}[|\mathbf{v}|_2^2] = \mathbb{E}[v_x^2 + v_y^2 + v_z^2] = 3\mathbb{E}[v_x^2]$ , given the fact that  $v_x$ ,  $v_y$  and  $v_z$  all have the same second moment. On finer meshes, the velocity magnitude is more accurate, but the mesh size is small enough to keep the cell Péclet number small.

What we do observe in numerical simulations is that the FEM solution undershoots near the boundary, attaining negative concentration values. This is a known phenomenon in the literature and it does not depend on the velocity field, but it is typical of diffusion problems with Dirichlet-type boundary conditions [167]. We solve

---

<sup>3</sup>For instance, in all the bias convergence plots of Chapter 4 the MC-FEM approximation to  $\mathbb{E}[||u^2||_{L^2}^2]$  is monotonically converging to the exact value from below.

this problem by resorting to mass-lumping, which is known to reduce this effect [167]. This ill-behaviour disappears as the mesh is refined to the extent that no undershootings are observed on the mesh used for point 1 and on the finer levels of the MLMC hierarchy. Note that in the MLMC case, it is immaterial whether non-physical behaviour is observed on the coarse levels as long as they still act as a good control variate for the finer levels.

## Solver and software

For the computations, we use a combination of the University of Oxford Mathematical Institute computing servers and the Abel supercomputing cluster<sup>4</sup>. We employ the FEniCS FEM software [120] and we solve the resulting linear system by using the PETSc [15] implementation of the GMRES algorithm preconditioned with the BoomerAMG algebraic multigrid algorithm from Hypre [57]. For the Matérn field sampling solver, we declare convergence when the absolute size of the preconditioned residual norm is below a tolerance of  $10^{-8}$ . For point 1, we estimate the PDFs of the output functionals with the kernel density estimation routines from the scikit-learn software [139].

## Further details

To address point 1, we solve models D1-2 and V1-3 by taking  $N = 3200$  MC samples. This ensures that, for all output functionals considered, we have a root mean square error below 1% of the expected value. We terminate the simulations after  $t$  has reached the value of  $T_1(\omega)$ , given by the maximum between 1 day and the maximum between the two activation times, cf. (6.9) and (6.10). In the model V3 case, these activation times might never occur and we just terminate the simulation after 1 day as soon as the total concentration in the white matter starts decreasing<sup>5</sup>. Convergence of the numerical solver for point 1 was verified with a convergence test comparing different mesh refinements and time steps for a set of deterministic worst-case models (with large velocities and small diffusion coefficients). Results are shown in the supplementary material of [44], but they are not reported in this thesis.

To address point 2, we only consider the more complex problems RF1 and RF2, but we do not consider the approximation of the PDFs of the output functionals and we do not compute the activation maps (6.9) and (6.10) as their lack of smoothness is likely to affect both MLMC and QMC convergence. MLMC techniques for density estimation are available [25, 70, 107], but they are not necessary here as the main objective is to compare the different Monte Carlo methods. Since we are not considering the activation maps, we always terminate the simulations after  $T_2 = 1$  day.

Some extra care must be taken in the QMC case: for QMC we consider the mesh of the coarsest MLMC level and we assume that the number of QMC samples needed is

---

<sup>4</sup>Owned by the University of Oslo and Uninett/Sigma2, and operated by the Department for Research Computing at USIT, the University of Oslo IT-department <http://www.hpc.uio.no/>. This research is supported by the NOTUR grant NN9316K.

<sup>5</sup>In this case we observe empirically that the white matter concentration starts to monotonically decrease to 0 because of the sink term in equation (6.1).

the same across levels, as suggested by Figure 5.6. We take a fixed Haar level  $\mathcal{L} = 4$  (Haar cell size roughly the same as the average mesh size). Since we have up to  $3 + 1$  input random fields (velocity + diffusivity), each of QMC dimensions  $N_f = 272$  (cf. Remark 5.6), and 1 input random variable, enforcing a correct variable ordering becomes complicated. Let  $n_v$  and  $n_f$  be the number of input random variables and fields of the problem considered. We sample a  $n_v + n_f N_f$  dimensional Sobol' sequence and we take the first  $n_v$  dimensions to correspond to the random variables. We then assign the  $(n_f k + i)$ -th dimension of the remaining  $n_f N_f$  dimensions to the  $k$ -th QMC dimension of the  $i$ -th field for  $i = 1, \dots, n_f$  and  $k = 0, \dots, N_f - 1$ . This way the remaining Sobol' point dimensions are spread equally among the input fields and a correct variable ordering is enforced<sup>6</sup>.

## 6.4 Clinical results: standard Monte Carlo simulations

### 6.4.1 Numerical results

We now present the results obtained by using standard Monte Carlo to quantify the uncertainty in the predictions of 5 different models (D1, D2, V1, V2, V3). Since the models are many, we now just show the numerical results obtained and we defer their discussion until the next subsection.

#### Deterministic diffusion as a baseline for parenchymal solute transport

To establish a baseline for the study, we first simulate the evolution of a tracer spreading in the SAS and in the parenchyma via diffusion only, using a deterministic and constant effective diffusion coefficient for gadobutrol ( $D^* = 1.2 \times 10^{-10} \text{ m}^2/\text{s}$ ). The resulting parenchymal tracer spread over 24 hours is shown in Figure 6.4. The tracer concentration increases first in inferior regions and in the gray matter, but does not penetrate deep into white matter regions within this time frame. In the sagittal plane (top), tracer concentration is higher as the cutting plane chosen is close to the fissure between the two brain hemispheres. This fissure is filled by free-flowing CSF in which the tracer spreads faster.

Figure 6.5a shows the tracer concentration on the boundary of the domain (concentration in the SAS) over time at the levels of three brain regions of interests (foramen magnum,  $z = -0.1 \text{ m}$ ; sylvian fissure,  $z = 0 \text{ m}$ ; and precentral sulcus,  $z = 0.1 \text{ m}$ ). During the first few hours, the boundary tracer concentration in the lower parts of the brain (foramen magnum) increases and peaks rapidly, while boundary tracer concentrations close to the other regions are lower, and with a longer time to peak. This is expected since we are simulating the injection of tracer in the spinal canal and tracer is therefore spreading upwards.

---

<sup>6</sup>See e.g. pg. 13-14 of <http://people.maths.ox.ac.uk/~giles/mc/mc/lec13.pdf> for a similar ordering strategy applied to multi-dimensional Brownian motion.

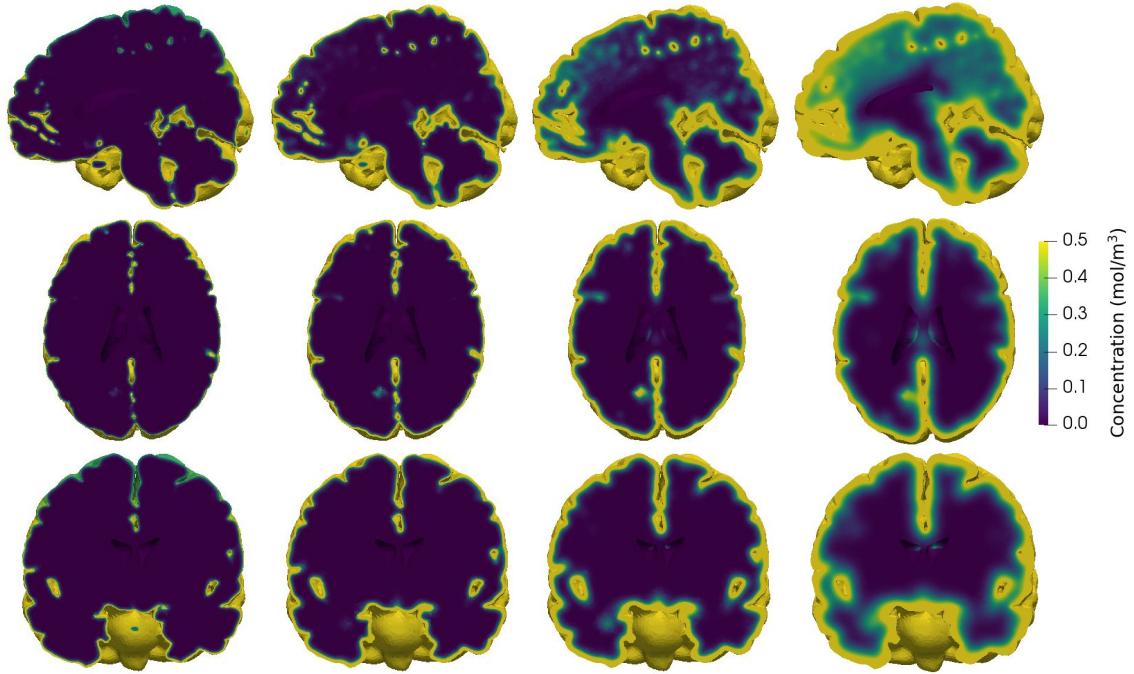


Figure 6.4: Tracer concentration in the parenchyma after (from left to right) 1, 3, 8 and 24 hours of diffusion in (from top to bottom) sagittal, transverse and coronal planes. Initially, most of the tracer is found in the lower regions. After 24 hours, tracer has penetrated substantially into the gray matter, but not into the deep, central regions.

In Figure 6.5b, concentration profiles are shown for three points within the parenchyma at different distances from the brain surface. The points are chosen along a line in the transverse plane ( $z = 0$ ) going from the outer surface to the ventricles. The tracer concentration at these points steadily increases after a few hours. This happens faster in regions closer to the SAS (green line). After 24 hours, the tracer concentration in all three points is still increasing, indicating that it is still not at steady-state.

### Quantifying the effect of uncertainty in effective diffusion magnitude

We first aim to quantify the effect of uncertainty in the magnitude of the effective diffusion coefficient on the time evolution of tracer in the gray and white matter. In particular, we now compute the tracer concentration and the other output functionals in the absence of a velocity field ( $\mathbf{v} = \mathbf{0}$ ), but with a gamma-distributed random variable as the diffusion coefficient (model D1).

The amount of tracer found in the gray and white matter differs both in magnitude and variability (Figure 6.6a-c). The expected amount of tracer in the gray matter increases rapidly and plateaus after approximately 15 hours, while the white matter does not reach steady state within the 24 hours of the simulation. There is little variability in the amount of tracer in gray matter throughout the 24 hour time span.

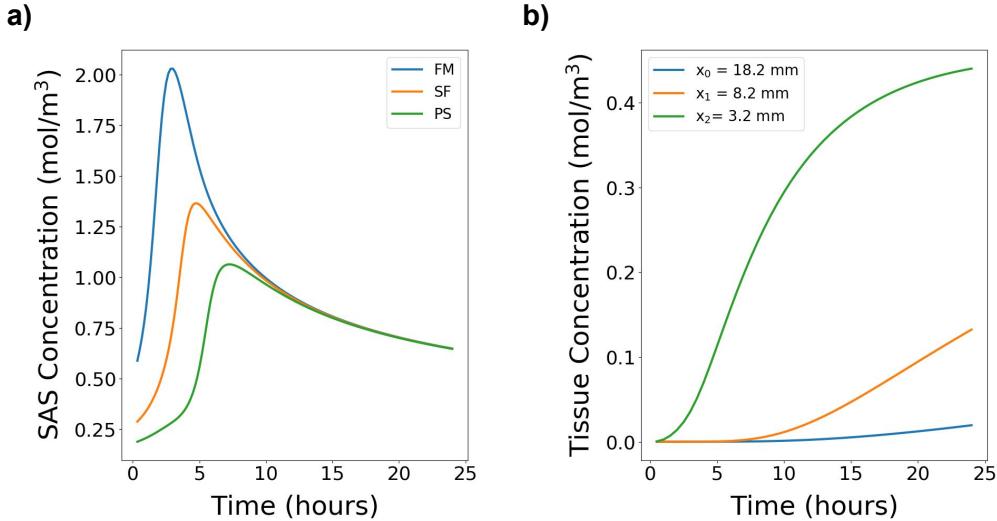


Figure 6.5: Tracer concentrations. (a) Tracer concentration in the SAS boundary ( $c_{\text{CSF}}$ ) at the level of the lower brain regions (foramen magnum, FM), middle regions (sylvian fissure, SF) and upper regions (precentral sulcus, PS). In the lower SAS, tracer concentration peaks at around 3 hours, while at the upper levels, peak concentration occurs later. Following peak values, the concentration in the SAS decreases as tracer enters the parenchyma. The SAS concentration on the boundary is modeled by (6.4) (b) Tracer concentration over time in three different points at a given distance from the brain surface. The points were chosen along a line lying on the transverse plane ( $z = 0$ ) going from the brain boundary to the ventricles.

We note that the prediction intervals shrink in Figure 6.6a as time increases, which matches the time evolution of the PDFs of the total amount of tracer (Figure 6.6c): as time evolves, the PDFs start concentrating over one value. This is because the concentration reaches the same constant steady state independently from the choice of diffusion coefficient. The PDFs become left skewed because the concentration approaches, but never surpasses the steady state value.

The amount of tracer in the white matter is initially almost constant, but starts increasing after 3 hours. The 99.73% prediction intervals are much larger in this case suggesting that for this model the value of the diffusivity strongly affects how much tracer can penetrate in the deeper regions of the brain.

The estimated PDF and CDF for the white matter activation time (i.e. time for 10% of tracer to reach the white matter) is shown in Figure 6.6d. We observe that the most likely white matter activation time is approximately 14 hours although there is quite some variability. Ten percent of the total injected tracer is therefore likely to penetrate within the white matter regions between 9.5 and 25 hours.

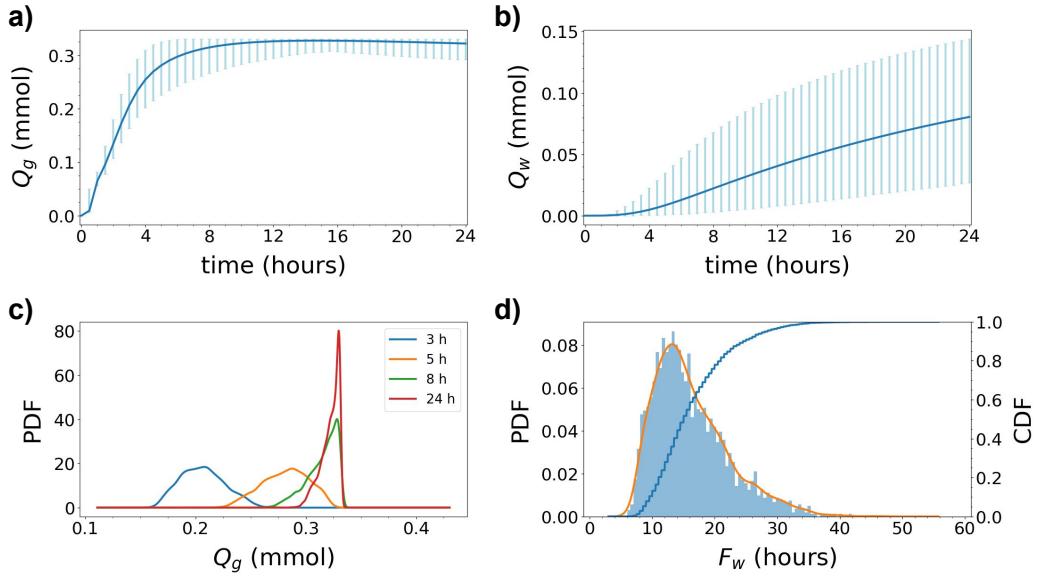


Figure 6.6: *Model D1 results.* Total tracer concentration in the (a) gray matter  $Q_g$  and (b) white matter  $Q_w$  over time (cf. equation (6.7)). The blue curves show the expected value, while the light blue vertical bars indicate 99.73% prediction intervals for the expectation computed using the estimated PDFs of the output functionals. (c) The estimated PDFs of  $Q_g$  at 3, 5, 8 and 24 hours after tracer injection. (d) Histogram of white matter activation time  $F_w$  (cf. equation (6.9)) (bars), corresponding estimated PDF using kernel density estimation (orange curve), and corresponding CDF. Uncertainty in the magnitude of the effective diffusion coefficients substantially impacts the amount of tracer found in the gray and white matter and the white matter activation time.

### Quantifying the effect of uncertainty in heterogeneous diffusivity

Brain tissue is heterogeneous [170], varies from individual to individual, and is clearly not accurately represented by a single diffusion value. To further investigate the effect of uncertainty in the diffusion coefficient and in particular to study the effect of spatial heterogeneity, we now consider the diffusion coefficient to be a spatially-varying random field (model D2).

The expected tracer concentration found in gray and white matter for model D2 are nearly identical to those resulting from model D1 (data shown later cf. Figure 6.9), but with substantially less variability and smaller prediction intervals.

However, when we turn to look at the tracer amount in the two smaller regions of interest (cf. (6.8)), variability in model D2 increases drastically (Figure 6.7). This is, perhaps, expected: each random field sample is a.s. comprised of many small regions (of size proportional to the correlation length) which are almost independent from each other. Consequently, the spatial average of the field across the whole brain can almost be seen as a “statistical” average between these regions, for which the variance

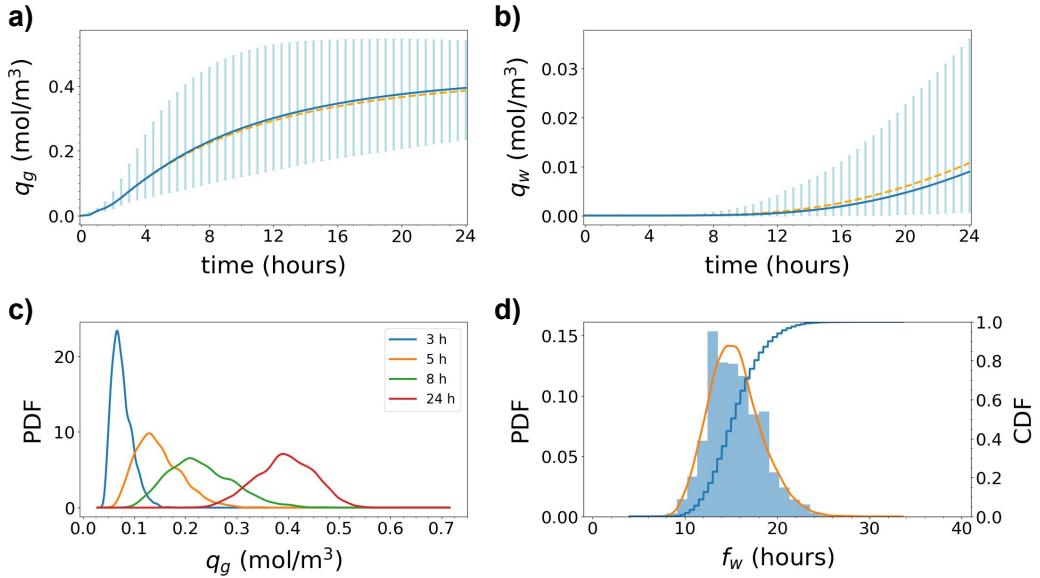


Figure 6.7: *Model D2 results.* The average tracer concentration in a small region within the (a) gray matter  $q_g$  and (b) white matter  $q_w$  as defined by (6.8). The blue curves show the expected value. The light blue vertical bars indicate the 99.73% prediction intervals. The dashed orange lines in (a) and (b) indicate the analogous expected value curve resulting from model D1 (constant diffusion only), for comparison. (c) The PDFs corresponding to  $q_g$  at 3, 5, 8 and 24 hours after tracer injection. (d) Histogram of white subregion activation time  $f_w$  as defined by (6.10) (bars), corresponding estimated PDF (orange curve), and CDF.

decreases as the number of regions grows. This means that smaller correlation lengths result in smaller regions and therefore smaller variability after integration. If we only average the field across a small region, the variability will instead be larger.

In the gray and white matter regions (Figure 6.7a), the expected average tracer concentration increases steadily with time. The expected average tracer concentration in the white matter is much smaller than in the gray matter (Figure 6.7b) and starts increasing much later, remaining close to zero after 24 hours. For the white region activation time, we observe reduced variability with respect to the model D1 case (Figure 6.7d), but a comparable mean.

### Quantifying the effect of glymphatic circulation

We now investigate the uncertainty surrounding ISF/CSF flow in paravascular spaces and the potential ISF flow in extracellular spaces, by considering a non-zero random velocity field. To investigate the effect of uncertainty in a glymphatic velocity model, we defined a random velocity field with correlation length corresponding to the typical distance between parenchymal arterioles and venules (model V1).

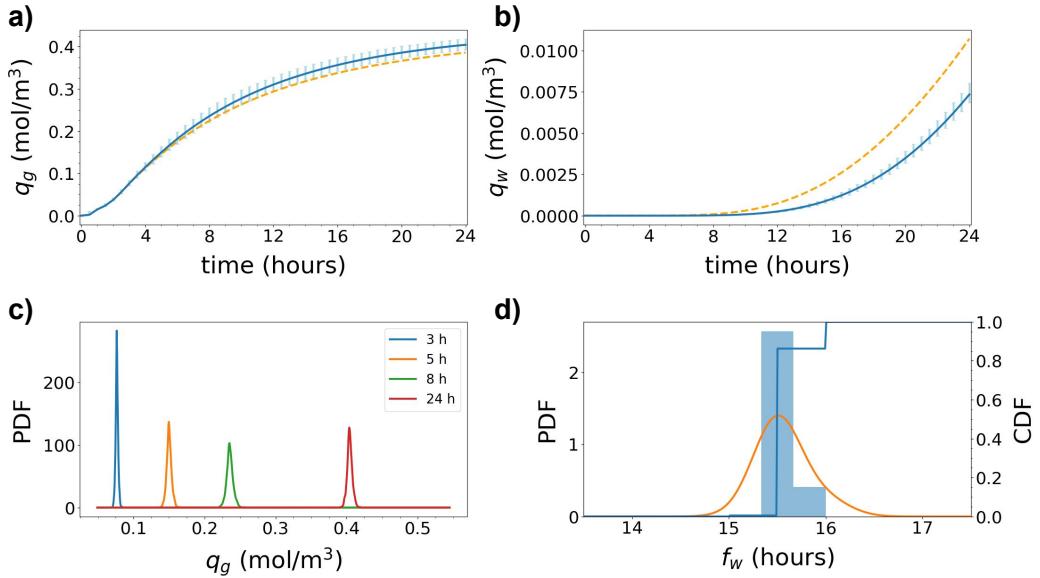


Figure 6.8: *Model V1 results.* The average tracer concentration in a small region within the (a) gray matter  $q_g$  and (b) white matter  $q_w$  as defined by (6.8). The blue curves show the expected value. The light blue vertical bars indicate the 99.73% prediction intervals. The dashed orange lines in (a) and (b) indicate the analogous expected value curve resulting from model D1 (constant diffusion only), for comparison. Expected values for  $q_g$  are nearly identical as for model D1 and D2, but variability is much lower. Expected values for  $q_w$  are lower than for model D1 and variability is much lower (c) The PDFs corresponding to  $q_g$  at 3, 5, 8 and 24 hours after tracer injection. The PDFs all have a very narrow support. (d) Histogram of white sub-region activation time  $f_w$  as defined by (6.10) (bars), and corresponding estimated PDF (orange curve) and CDF.

The expected amounts of tracer found in the whole gray and whole white matter for model V1 are nearly identical to those found for model D2 and model D1, while the variability is minimal (data shown later cf. Figure 6.9). Thus, on average, small random variations in fluid velocity did not increase (or decrease) the tracer distribution into the parenchyma on a global scale.

We observe that the expected  $q_g$  increases marginally faster in the V1 glymphatic velocity model than for pure diffusion, cf. Figure 6.8a-b: at 24 hours,  $q_g$  is 2.5% higher for V1 than for D1. On the other hand, the expected  $q_w$  increases faster with pure diffusion: at 24 hours,  $q_w$  is 34% lower for V1 than for D1.

The variability in both gray and white local space-average tracer concentration is extremely small (cf. Figures 6.8a-c), and the activation time  $f_w$  shows low variability (figure 6.8d) as well (the activation time of all samples lies between 15 and 16 hours). This shows that uncertainty in the velocity field has a much smaller impact on the solution uncertainty than in the random diffusion case.

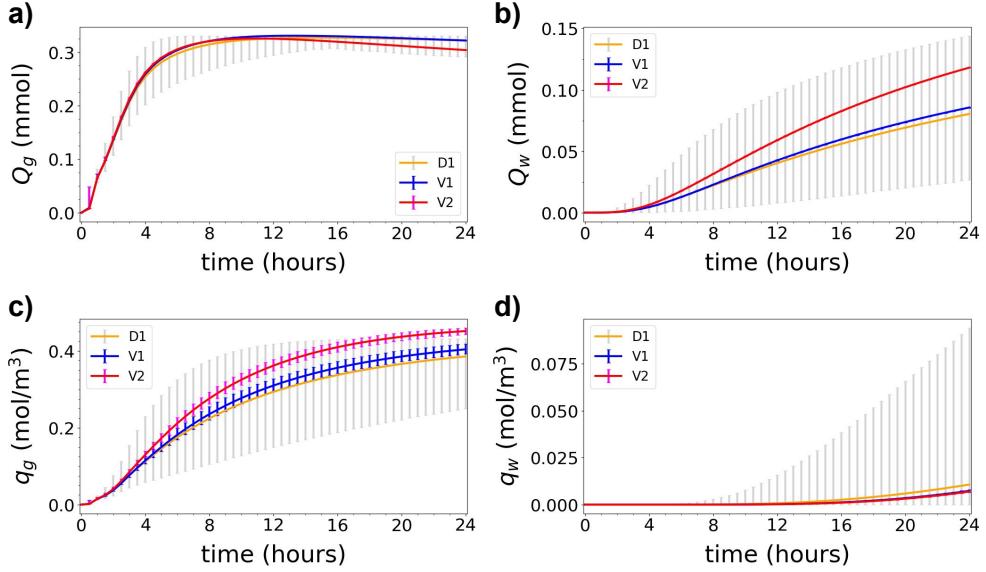


Figure 6.9: *Model V2 results. Model V2 (red) in comparison with models D1 (orange) and V1 (blue). The total amount of tracer in the (a) gray matter  $Q_g$  and (b) white matter  $Q_w$ , as defined by (6.7), over time. The average tracer concentration in a subregion of (c) gray matter  $q_g$  and (d) white matter  $q_w$ , as defined by (6.8), over time. The curves show the expected values while vertical bars indicate the 99.73% prediction intervals of the different models.*

### Quantifying the effect of glymphatic directionality

We now investigate the hypothesis that glymphatic circulation has a directionality given by large vascular structures [104, 143]. To assess whether and how such a directionality affects the tracer distribution, we remove the assumption that the velocity field has zero-mean and we add a net flow field to the velocity field representing the directionality in the glymphatic circulation (model V2).

With more fluid entering the brain from below, as illustrated by the streamlines of Figure 6.2c, the total amount of tracer in the brain increases. For the expected amount of tracer in gray matter, however, model V2 was in very good agreement with models D1 and V1 (Figure 6.9a). After 13 hours, the amount of tracer found in the gray matter is higher for model D1 than for model V2. In model V2, more of the tracer is found deeper in the gray matter and eventually moves to the white matter. We note that the uncertainty associated with the velocity fields barely propagates to the amount of tracer in the gray and white matter (Figure 6.9a-b).

The expected amount of tracer in the white matter  $Q_w$  increases substantially by the introduction of the directional velocity field (Figure 6.9b). The expected value curve starts deviating from the other models after 4-5 hours, and the difference increases with time. At 24 hours, the expected amount of tracer found in the white matter  $Q_w$  is 50% larger for model V1 as for model D1, although still within the

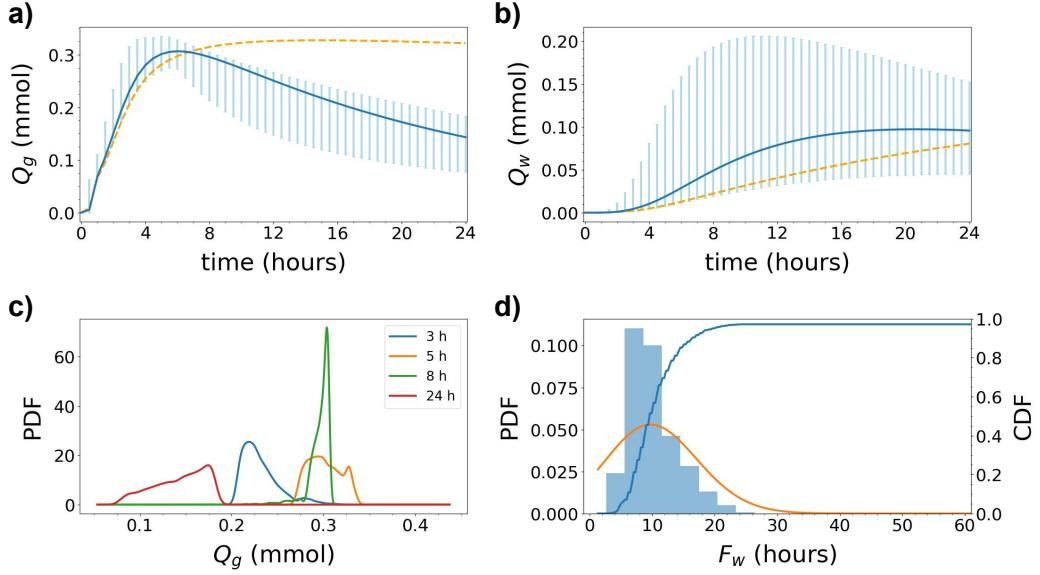


Figure 6.10: *Model V3 results.* The total amount of tracer in the (a) gray matter  $Q_g$  and (b) white matter  $Q_w$  over time;  $Q_g$  and  $Q_w$  as defined by (6.7). The blue curves show the expected value. The light blue vertical bars indicate the 99.73% prediction intervals. The dashed orange lines in (a) and (b) indicate the analogous expected value curve resulting from model D1 (constant diffusion only), shown for comparison. Large variations in the white matter is found depending on the inflow velocity. (c) The estimated PDFs corresponding to  $Q_g$  at 3, 5, 8 and 24 hours after tracer injection. (d) Histogram of white matter activation time  $F_w$  as defined by (6.9) (bars), corresponding estimated PDF (orange curve), and CDF. We note that the plotted CDF does not reach 1 since some samples (roughly the 3%) never reach the white region activation threshold.

99.73% prediction interval. The directional velocity field also induces an increase in the expected average tracer concentration in the gray subregion  $q_g$  (Figure 6.9c), while for  $q_w$  model V1 and V2 are in close agreement (Figure 6.9d).

### Quantifying the effect of paraarterial influx with drainage

A number of open questions remain in the context of glymphatic and paravascular outflux routes. To further investigate potential ISF outflux pathways, we finally consider a model representing paraarterial influx combined with ISF drainage (model V3).

In model V3 (paraarterial inflow with drainage) the amount of tracer found in the parenchyma for the early time points is larger than in the other model (Figure 6.10). However, in this case, the concentration reaches a peak and then starts to decrease monotonically because of the sink term.

For the white matter, the expected amount of tracer increases with time faster than with pure diffusion, and seems to peak at approximately at around 20 hours before slowly decreasing. The level of uncertainty in the white matter concentration is large, and the expected value is larger than in the other models. However, after 24 hours the concentration values predicted by both models seem to agree.

The white matter activation time is, on average, lower for model V3 compared to previous models (Figure 6.10d), but there is substantial variability as in some cases the activation time is never reached (in 3% of the computed samples).

### 6.4.2 Discussion

We now compare our simulations with the healthy patient MRI study by Ringstad et al. [146]. We remark that the MRI-study only provides quantitative values of “signal intensity” and that the map between signal intensity and tracer concentration is typically a non-linear function (unknown in this case) [40]. We are thus unable to make a direct comparison between these two quantities. However, we can still assume that a peak in signal intensity corresponds to a peak in tracer concentration and that a comparison of the time to peak between the model results and experiments is still possible.

*Remark 6.2.* Since obtaining quantitative concentration measurements via MRI is extremely difficult, adopting a Bayesian approach to estimate the input parameters given concentration measurements is extremely complicated and thus the only feasible approach for the moment is the UQ forward problem.

Similarly as in the MRI study by Ringstad et al. [146], the tracer first spreads to the lower regions of the brain that are closer to the tracer injection site and does not reach the ventricles. In models with diffusion only, the amount of tracer in the gray matter peaks at approximately 15 hours, while in the MRI study, the time to peak in selected regions of interest is between 12 and 24 hours [146]. However, unlike in the clinical experiments, white matter concentration does not peak within 24 hours of simulation in models D1 and D2 and the predicted uncertainty in the white matter concentration cannot explain this discrepancy. This suggests that a mechanism other than diffusion is in action to transport tracer into deeper regions of the brain.

In our glymphatic circulation model (model V1), the stochastic velocity field does not increase tracer distribution to the brain on average. However, when we focus our attention to the smaller regions, tracer concentration may increase by up to 13% compared to diffusion alone. These results compare favourably with a clinical study by Iliff et al. [94] and can mathematically be expected: the effect of a zero-mean velocity field ( $\mathbb{E}[\mathbf{v}] = \mathbf{0}$ ) is averaged out when integrating the concentration over regions of size much larger than the field correlation length  $\lambda$ , yielding results which are comparable to the pure diffusion case ( $\mathbf{v} = \mathbf{0}$ ). However, this averaging effect does not occur when looking at a region of size proportional to or smaller than  $\lambda$ .

*Remark 6.3.* We recall that the choice of a zero-mean field comes from the assumption (assumption number 2 in model V1) that the blood vessel structure is independent from the location within the parenchyma. Let us now re-interpret each Monte Carlo

sample as one measurement from a patient. These results then suggest that clinical experiments willing to investigate the presence of ISF bulk flow across the brain must be designed so that measurements are taken in regions where the vasculature is consistent across patients, e.g. always close to an artery rather than at random locations. This way the effect of a convective velocity will not be averaged out when taking the mean between the measurements.

When modelling paraarterial influx combined with drainage (model V3), the time to peak is reduced to 6-8 hours in the gray matter. Although we lack quantitative information about the drainage parameter, we observe that drainage/clearance would reduce both the time to peak and relative tracer concentration in the brain compared to diffusion alone. In the glymphatic directionality model (model V2), the presence of a paravascular directional velocity also decreases the expected time to peak tracer concentration in gray matter down to 11 hours (compared to 15 hours for pure diffusion). Thus, when experimental data suggests a time to peak shorter than for diffusion alone, it is not clear whether this is due to increased glymphatic function or increased clearance by parenchymal drainage.

In our models, the white matter (and the small regions) is where the effect of a convective velocity becomes most prominent. The only model modification causing an expected time to peak in white matter of approximately 24 hours is with a paraarterial inflow and drainage (model V3). In this model, the upper limit of the 99.73% prediction interval peaks at approximately 12 hours, which is more comparable to the rapid increase in tracer concentration observed in the white matter of healthy subjects [147].

In our experiments it is evident that uncertainty in the diffusion parameter significantly affects the model predictions, while uncertainty in the velocity field does not seem to substantially propagate to the output. Furthermore, even if the expected concentrations of models D1 and D2 appear to be the same, the uncertainty in the concentration values is lower for D2 (random field) than for D1 (random variable). This suggests that: 1) having an accurate estimate of tracer diffusivity is essential to be able to trust what is predicted by deterministic models, and 2) the tracer distribution to large brain regions can still be well approximated using a constant diffusion value, provided that the diffusion coefficient is estimated accurately. If accurate measurements are not possible, considering a heterogeneous diffusion coefficient estimated with multiple measurements in different brain regions can significantly help to reduce the uncertainty.

Although diffusion may act as the main transport mechanism in the parenchyma [92, 159], it is also clear that diffusion alone is not sufficient to transport tracer deep into the parenchyma and that net bulk flow (i.e. advection) is also needed to enhance tracer penetration (as in models V2 or V3). Our results therefore support the hypothesis that solutes within the ISF spread due to a combination of diffusion and convection.

### 6.4.3 Section conclusions

In this study we investigated the variability in brain tracer concentration resulting from uncertainty in the diffusion and convection parameters. We designed five computational models, each representing different diffusion and convection regimes and different assumptions about ISF fluid movement and we used standard Monte Carlo to compute the probability distributions of all output functionals of interest.

The results from this study show that uncertainty in the diffusion parameters substantially impact the amount of tracer in gray and white matter, and the average tracer concentration in small regions within the gray and white matter. Consequently, multiple or more accurate measurements of the diffusion coefficient are needed to reduce the uncertainty in model predictions. Nevertheless, even with moderate uncertainty in the diffusion coefficient and a resulting four-fold variation in white matter tracer concentration, discrepancies between simulations of pure diffusion and experimental data are too large to be attributed to uncertainties in the diffusion coefficient alone.

A convective velocity field, representing the glymphatic circulation, increases tracer concentration in the brain as compared to pure diffusion. However, this increased concentration depends on a directional structure of the velocity field and more accurate ISF flow measurements would be needed to establish whether such a flow structure is effectively present. Overall, more quantitative experimental data are needed to obtain a deeper understanding of the brain glymphatic function.

## 6.5 Computational comparison of the methods presented in this thesis

### 6.5.1 Numerical results

We now compare the efficiency of standard MC, QMC and MLMC when employed to solve models RF1 and RF2. In what follows, we let  $\mathcal{T} = \{30k \min\}_{k=1}^{k=48}$  and define

$$\mathcal{Q} = \{Q_g(t), t \in \mathcal{T}\} \cup \{Q_w(t), t \in \mathcal{T}\} \cup \{q_g(t), t \in \mathcal{T}\} \cup \{q_w(t), t \in \mathcal{T}\}, \quad (6.24)$$

to be the set of all the output functionals of interest considered (cf (6.7) and (6.8)), with the exception of the activation maps (6.9) and (6.10).

To begin with, we estimate the MLMC parameters  $\alpha$ ,  $\beta$  and  $\gamma$  of Theorem 2.1. Since we are considering the estimation of multiple output functionals, we estimate  $\alpha$  and  $\beta$  by monitoring the quantities

$$\max_{Q \in \mathcal{Q}} |\mathbb{E}[Q_\ell - Q_{\ell-1}]|, \quad \max_{Q \in \mathcal{Q}} \mathbb{V}[Q_\ell - Q_{\ell-1}]. \quad (6.25)$$

We expect  $\alpha = 2$ ,  $\beta = 2\alpha = 4$  and  $\gamma = 4$  since we have a second-order method in both time and space in this case and we are using a multigrid-preconditioned Krylov method (cf. Section 6.3). To estimate the bias and variance in (6.25) we take  $N = 4000$  samples on the first two levels and  $N = 100$  on the finest level. The

number of vertices on the finest level of the MLMC hierarchy is quite large (22 282 705 vertices), resulting in a memory burden of  $\approx 50$  GB to just load the mesh, the box mesh in which the brain mesh is embedded (cf. Section 4.2.3) and the FEM subspaces required. Additionally, solving one instance of equation (6.1) on this mesh takes more than 24 hours in serial. We thus compute 100 samples for each model on the finest level.

In Figures 6.11 and 6.12 we show the convergence behaviour of the (logarithm of the) quantities in (6.25) for model RF1 and RF2 respectively. We see that almost all estimated convergence orders match with our predictions apart from the bias of model RF1 that decays more rapidly than expected. In this case, we are likely observing a pre-asymptotic regime and the higher-than-expected bias convergence rate seems to be decaying as  $\ell$  increases. The estimated variance convergence order for both models is around  $\hat{\beta} \approx 4.15$ , which is just slightly above the theoretical value. When estimating  $\gamma$  by taking the average wall-clock time of each sample as a proxy, we obtain  $\hat{\gamma} \approx 4.09$ , which is close to the  $\gamma = 4$  expected.

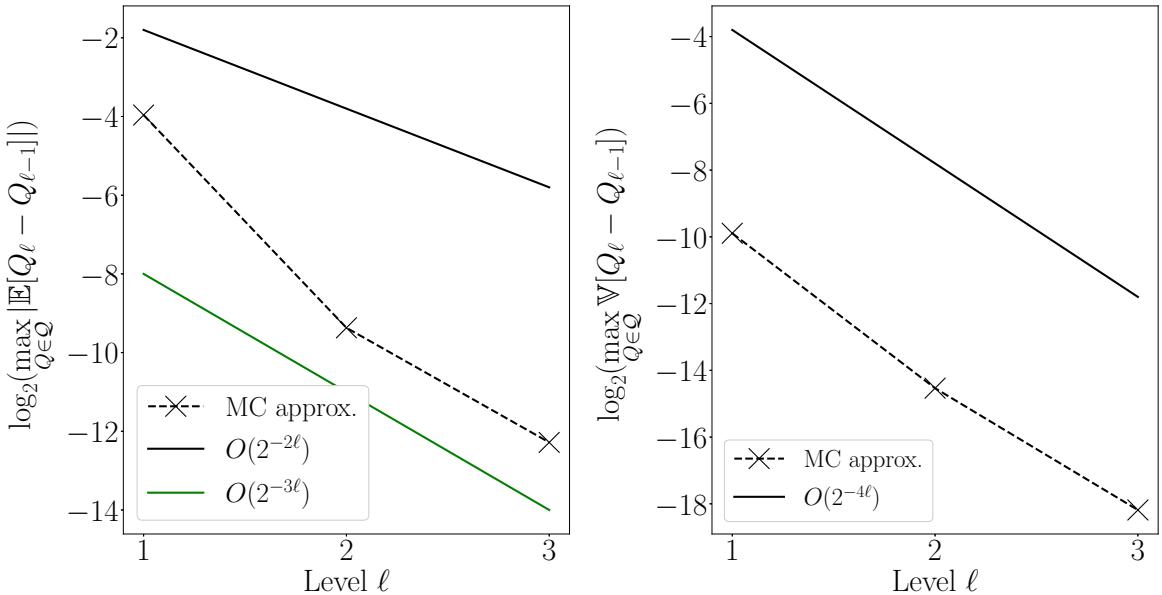


Figure 6.11: *Convergence behaviour of the FEM approximation to the solution of model RF1. The estimated convergence order for the variance agrees with our predictions and with what expected by the theory in the diffusion-only case [166]. The bias convergence order observed is instead higher than expected. Estimated parameters via linear regression:  $\hat{\alpha} \approx 4.16$ ,  $\hat{\beta} \approx 4.15$ .*

Note that we have a finite number of meshes available and consequently the version of MLMC considered here is “weaker” than the true MLMC algorithm presented in Section 2.2. In fact, we are unable to reduce the bias below the threshold given by the finest mesh of the hierarchy without resorting to more advanced techniques (cf. Remark 6.4). However, we can still balance the relative weight of bias and statistical error by choosing two different values of the MLMC weight parameter  $\theta$  introduced by Haji-Ali et al. in [82] (cf. Section 2.2). The mean square error for an output functional

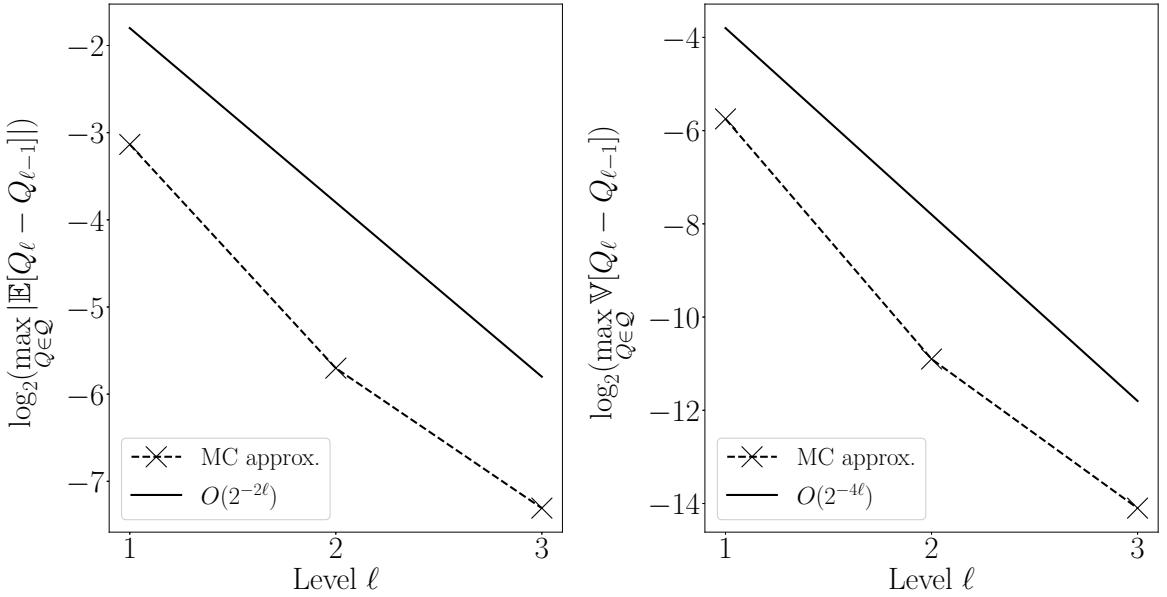


Figure 6.12: *Convergence behaviour of the FEM approximation to the solution of model RF2. The estimated convergence orders agree with our predictions and with what expected by the theory in the diffusion-only case [166]. Estimated parameters via linear regression:  $\hat{\alpha} \approx 2.09$ ,  $\hat{\beta} \approx 4.18$ .*

$Q$  is given by

$$MSE = \hat{V} + \mathbb{E}[\hat{Q} - Q]^2, \quad (6.26)$$

where  $\hat{Q}$  is the MLMC estimator for  $\mathbb{E}[Q]$  and  $\hat{V}$  is the estimator variance. To ensure a MSE below  $\varepsilon^2$ , we therefore impose

$$\hat{V} \leq (1 - \theta)\varepsilon^2, \quad \mathbb{E}[\hat{Q} - Q]^2 \leq \theta\varepsilon^2, \quad (6.27)$$

with  $\theta \in (0, 1)$  weighting the two terms. Small values of  $\theta$  reduce the number of samples needed and are therefore preferred when the bias is small. Conversely, large values are beneficial when the bias is large as they allow to achieve smaller tolerances without adding finer levels to the hierarchy. Here, we are unable to take more than 100 samples on the finest level, which further restricts the amount of effort we can dedicate to reducing the statistical error. For this reason, we choose the largest values of  $\theta$  that do not make the number of samples on the finest level exceed 100. Numerical experimentation yields the values  $\theta = 0.04$  for model RF1 and  $\theta = 0.7$  for RF2. Note that in the model RF1 case, the bias is much smaller (compare Figures 6.11 and 6.12), hence why the chosen  $\theta$  is smaller as well. In the standard MC and QMC case, we also require that  $\hat{V} \leq (1 - \theta)\varepsilon^2$ , where  $\hat{V}$  in this case is the estimator of either method.

*Remark 6.4.* In practice, it is possible to reduce the MLMC estimator bias by augmenting MLMC with Richardson-Romberg extrapolation [69, 113]. However, we leave this enhancement for future research.

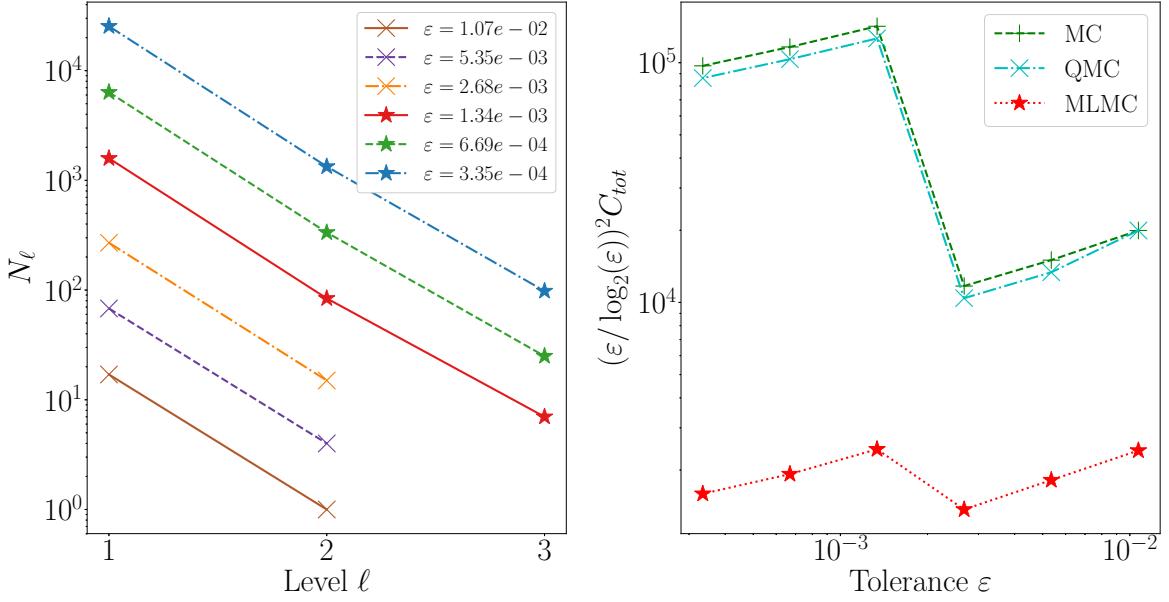


Figure 6.13: Convergence of standard MC, QMC and MLMC for the solution of model RF1 ( $\theta = 0.04$ ). In the plot on the left we show how the MLMC algorithm automatically selects the optimal number of samples  $N_\ell$  on each level to achieve a given tolerance  $\varepsilon$ . In the plot on the right we compare the efficiency of the methods for different tolerances. The savings of MLMC with respect to standard MC and QMC are considerable, while QMC barely improves on standard MC (see main text).

In the plots on the left in Figures 6.13 and 6.14, we show the optimal number of samples chosen automatically by MLMC on each level as the root mean square error tolerance  $\varepsilon$  is reduced. The maximum level chosen is increased as  $\varepsilon$  decreases in order to satisfy the bias tolerance. Note that the smallest values of  $\varepsilon$  considered correspond to the lowest bias tolerance that standard MLMC can achieve with an upper limit of 100 samples on the finest level (cf. Remark 6.4).

In the plots on the right in Figures 6.13 and 6.14 we compare the total computational cost  $C_{tot}$  of standard MC, MLMC and QMC for the solution of model RF1 and RF2 respectively. Since we are in the  $\beta = \gamma$  case, we expect an overall MLMC complexity of  $\varepsilon^{-2}(\log \varepsilon)^2$  for a root mean square error tolerance of  $\varepsilon$  (cf. Theorem 2.1). For this reason, we plot the quantity  $\varepsilon^2(\log \varepsilon)^{-2}C_{tot}$  on the  $y$ -axis as this should be  $O(1)$  in the MLMC case, giving a horizontal MLMC cost line. In fact, the MLMC cost lines shown in the figures slightly oscillate, but they are well-fitted by a horizontal line (estimated slope  $\approx -0.04$ , considering just the 4 smallest tolerances).

We note that for both models MLMC significantly outperforms both QMC and standard MC, with a  $O(100)$  factor of improvement with respect to standard MC. While the qualitative behaviour of standard MC and MLMC is consistent between the two models, QMC behaves differently: in the model RF2 case (Figure 6.14) we see QMC considerably improving with respect to standard MC. In the model RF1 case (Figure 6.13), instead, the improvement is negligible and QMC performs essentially the same as standard MC. This behaviour could, perhaps, be expected by looking

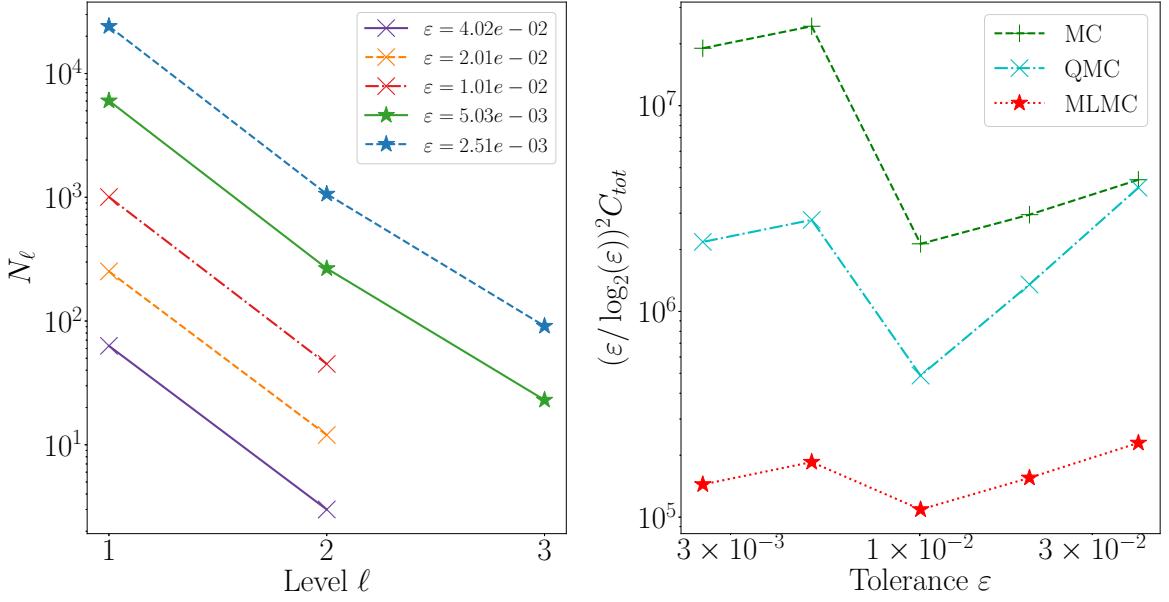


Figure 6.14: *Convergence of standard MC, QMC and MLMC for the solution of model RF2 ( $\theta = 0.7$ )*. In the plot on the left we show how the MLMC algorithm automatically selects the optimal number of samples  $N_\ell$  on each level to achieve a given tolerance  $\varepsilon$ . In the plot on the right we compare the efficiency of the methods for different tolerances. MLMC significantly outperforms QMC, which in turn considerably outperforms standard MC.

at the formulations of models RF1 and RF2 (cf. Table 6.1). While the stochastic input in model RF2 is comprised by 1 random field and 1 random variable, model RF1 depends on 4 random fields and 1 random variable. Given the performance observed in the figures, it is clear that this higher input dimensionality is affecting the good convergence properties of QMC, causing the convergence rate to decay to a  $O(N^{-1/2})$  standard MC rate. The fact that QMC performance degrades with high input dimensions is well-known [35] and was discussed in this thesis in Section 2.1. It therefore appears that the (ML)QMC method presented in Chapter 5 is not robust with respect to the number of random field inputs, at least in 3D where the dimensionality is larger<sup>7</sup>.

*Remark 6.5.* Adding a coarser level to the mesh hierarchy, given by the original Colin27 human adult brain atlas mesh [58] (version 2) did not improve the performance of MLMC.

*Remark 6.6.* In this section we did not consider MLQMC. From the results in Figure 6.13 it is however clear that no additional improvement can be achieved in the model RF1 case. On the other hand, the model RF2 case (cf. 6.14) seems to be amenable to MLQMC, but we leave this investigation for future research.

<sup>7</sup>We did not observe this ill-behaviour in analogous numerical tests performed on a convection-diffusion PDE with random coefficients on a square domain.

### 6.5.2 Section conclusions

In this section we have considered the solution of two models of interest with most of the sampling methods presented in the thesis (with the exception of MLQMC). Even under a restriction on the maximum number of samples on the finest level and on the finest mesh available MLMC significantly outperforms all other methods, yielding an improvement factor of roughly 100 with respect to standard MC. QMC outperforms standard MC by a factor of approximately 10 in the model RF2 case. However, in the model RF1 case QMC barely improves on standard MC, possibly because of the high number of input random fields in 3D, which make the problem extremely high-dimensional. Overall, for this application, MLMC achieves the best performance and should be preferred. Possibly MLQMC could bring additional computational gains for small numbers of input random fields.

# Chapter 7

## Conclusions and future research directions

### 7.1 Summary of results

In this thesis we focused on two main problems: 1) the efficient sampling of spatial white noise and Matérn field realisations within Monte Carlo methods, and 2) the quantification of uncertainty in simulations of brain tracer transport. We mainly considered the Whittle SPDE approach to Matérn field sampling, for which the sampling problem is recast as the numerical solution of an elliptic PDE driven by spatial white noise.

In Chapter 3, we presented a new derivation of an *a priori* error estimate for the finite element solution of the Whittle SPDE. At the time in which it was derived, this result improved with respect to the existing literature in multiple regards: the proof did not require the approximation of white noise in practice, and higher order elliptic operators and the  $p$ -refinement case were considered.

Within the SPDE approach, the sampling of white noise realisations can become the main computational bottleneck. In fact, when continuous Lagrange elements are used, the mass matrix is sparse, but typically not diagonal and computing its Cholesky factorisation has  $O(m^3)$  cost complexity. In Chapter 4 we presented a new algorithm for sampling white noise realisations with optimal linear complexity even in the case in which the FEM mass matrix is not diagonal. Furthermore, we showed how coupled white noise realisations can still be sampled with the same complexity even within a non-nested MLMC hierarchy. In this case, a supermesh construction is required and the overall sampling cost is linear in the number of supermesh cells. As an auxiliary result, we also proved that when the parent meshes are quasi-uniform, the number of cells of the supermesh is in turn linear in the number of cells of the parent meshes. This result is new, improves on previously known bounds, and is useful in the complexity analysis of other applications of supermeshes.

In Chapter 5 we extended this framework to the QMC and MLQMC case. We showed that by using a Haar wavelet expansion of white noise and a hybrid MC/QMC approach it is possible to obtain considerable computational gains in a pre-asymptotic

regime. Within this regime, which depends on the Haar level chosen, the convergence with respect to the number of samples is faster than  $O(N^{-1/2})$ . For this method, a two-way and a three-way supermesh is required in the QMC and MLQMC case respectively. However, the overall sampling complexity remains linear with respect to the number of supermesh cells which is in turn linear in the number of parent mesh cells.

Finally, in Chapter 6 we applied these new methods (with the exception of MLQMC) to the solution of a new PDE model with random coefficients. This model describes the movement of a tracer within the ISF and was solved on a real-life brain geometry derived via MRI in 3D. We first solved 5 different “clinical” instances of this model by using standard MC. Each of these reflects a different hypothesis proposed in the medical literature regarding ISF movement and brain solute clearance. We then applied standard MC, QMC and MLMC to assess which algorithm was the most efficient. MLMC resulted to be the most computationally advantageous, outperforming both QMC and standard MC.

The investigation on the 5 different “clinical” models revealed that the diffusion coefficient is the largest source of uncertainty in the PDE, showing that accurate diffusivity values are needed if deterministic models are to be trusted. Nevertheless, the discrepancies between the diffusion-only models and experimental data were too large to be solely attributed to diffusion uncertainty. In fact, our results suggest that the presence of a convective velocity field is needed to justify the presence of tracer deep into the brain white matter. To the author’s knowledge, this was the first time that advanced Monte Carlo methods have been used for UQ in brain modelling and the first UQ study ever performed on simulations of tracer transport within the brain.

## 7.2 Future research directions

In this thesis we only considered the SPDE approach to Matérn field sampling. However, many more methods are available in the literature, as discussed in Chapter 2. For each method, leading-order complexity estimates are available, but the proportionality constants are often difficult to estimate theoretically. A computational comparison aimed at assessing which method performs the best in which situation would extremely insightful. To the author’s knowledge, no such comparison has ever been performed in the literature.

Even though the numerical results of Chapter 5 support the efficacy of the (ML)QMC algorithm presented, only a partial convergence result was derived. It would be interesting to derive a full convergence theory for the proposed strategy, possibly by using randomly shifted lattice rules and extending the theory derived by Hermann and Schwab [88]. Such a result would yield formulas for the optimal number of samples and optimal Haar level on each MLQMC level, which in turn would increase the efficiency of MLQMC. In Chapter 6 we saw how the performance of QMC degraded when the number of input random fields was increased. A full convergence theory considering multiple inputs could improve our understanding of this behaviour.

From a more practical point of view, it would be interesting to extend our (ML)QMC

algorithm to general higher degree wavelets and domains (cf. remarks 5.18 and 5.19). The first enhancement (general wavelets) could possibly improve the convergence rate with respect to the number of samples, while the second (general domains) would reduce the supermeshing complexity and consequently the white noise sample cost. Another way to reduce the MLQMC sampling complexity is to implement a dedicated three-way supermeshing algorithm: the strategy we used (computing two two-way supermeshes) is likely to be suboptimal.

In this thesis, we mainly considered spatial white noise driven SPDE with integer-order operators. However, so as to be able to sample fields with the whole range of Matérn covariances, we must also be able to deal with the fractional case efficiently. There are three main approaches used in the literature: the first consists in rewriting the inverse fractional operator as a Dunford-Taylor contour integral in the complex plane whose integrand only depends on the inverse of an integer-order operator. This integral can then be approximated via exponentially convergent quadrature rules for which the number of quadrature nodes increases with the mesh size [27, 30, 42, 83]. The second approach uses the Caffarelli–Silvestre extension of the PDE from  $\mathbb{R}^d$  to  $\mathbb{R}^d \times [0, \infty)$  [34]. This yields a standard integer-order PDE, to be solved on the extended  $\mathbb{R}^d \times [0, \infty)$  domain, for which standard FEMs can be used. Finally, the third consists in employing a rational Krylov subspace method [33]. Using any of these techniques results in an increased sampling cost, which could however be mitigated by developing more advanced MLMC sampling techniques.

There is great need for more UQ in brain simulations and other models could be investigated. An option would be to consider the multiple-network poroelastic theory equations that simultaneously describe the movement of multiple brain fluids and the evolution of brain matter displacement [171]. Finally, given suitable experimental data, it would be extremely insightful to also solve the inverse UQ problem and obtain a better estimate of the physiological parameters of the brain of medical interest, such as the velocities of various brain fluids or brain matter permeability.

# Bibliography

- [1] N. J. Abbott. “Evidence for bulk flow of brain interstitial fluid: significance for physiology and pathology”. In: *Neurochemistry International* 45.4 (2004), pp. 545–552.
- [2] N. J. Abbott, M. E. Pizzo, J. E. Preston, D. Janigro, and R. G. Thorne. “The role of brain barriers in fluid movement in the CNS: is there a ‘glymphatic’ system?” In: *Acta Neuropathologica* (2018), pp. 1–21.
- [3] P. Abrahamsen. *A Review of Gaussian Random Fields and Correlation Functions*. 2nd ed. Norwegian Computing Center, 1997. ISBN: 8-25-390435-5.
- [4] D. L. Adams, V. Piserchia, J. R. Economides, and J. C. Horton. “Vascular supply of the cerebral cortex is specialized for cell layers but not columns”. In: *Cerebral Cortex* 25.10 (2015), pp. 3673–3681.
- [5] N. J. Albargothy, D. A. Johnston, M. MacGregor-Sharp, R. O. Weller, A. Verma, C. A. Hawkes, and R. O. Carare. “Convective influx/glymphatic system: tracers injected into the CSF enter and leave the brain along separate periarterial basement membrane pathways”. In: *Acta Neuropathologica* 136.1 (2018), pp. 139–152.
- [6] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells. “The FEniCS Project Version 1.5”. In: *Archive of Numerical Software* 3.100 (2015).
- [7] *Alzheimer’s Research UK*. 2017. URL: [www.dementiastatistics.org](http://www.dementiastatistics.org).
- [8] R. A. Armonda, C. M. Citrin, K. T. Foley, and R. G. Ellenbogen. “Quantitative cine-mode magnetic resonance imaging of Chiari I malformations: an analysis of cerebrospinal fluid dynamics”. In: *Neurosurgery* 35.2 (1994), pp. 214–224.
- [9] M. Asgari, D. De Zélicourt, and V. Kurtcuoglu. “Glymphatic solute transport does not require bulk flow”. In: *Scientific Reports* 6 (2016), p. 38635.
- [10] A. Aspelund, S. Antila, S. T. Proulx, T. V. Karlsen, S. Karaman, M. Detmar, H. Wiig, and K. Alitalo. “A dural lymphatic vascular system that drains brain interstitial fluid and macromolecules”. In: *Journal of Experimental Medicine* 212.7 (2015), pp. 991–999.
- [11] R. N. Auer. “Histopathology of brain tissue response to stroke and injury”. In: *Stroke*. 6th ed. Elsevier, 2016, pp. 47–59. ISBN: 978-0-3232-9544-4.
- [12] I. Babuška and M. Suri. “The  $p$  and  $h-p$  versions of the finite element method, basic principles and properties”. In: *SIAM Review* 36.4 (1994), pp. 578–632.

- [13] M. Bachmayr, I. G. Graham, V. K. Nguyen, and R. Scheichl. *Unified Analysis of Periodization-Based Sampling Methods for Matérn Covariances*. 2019. arXiv: 1905.13522.
- [14] E. N.T. P. Bakker, D. M. P. Naessens, and E. VanBavel. “Paravascular spaces: entry to or exit from the brain?” In: *Experimental Physiology* 104.7 (2019), pp. 1013–1017.
- [15] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. R. Brune, K Buschelman, V Eijkhout, W Gropp, D Kaushik, M. G. Knepley, and Others. *PETSc users manual revision 3.8*. Tech. rep. Argonne National Laboratory (ANL), 2017.
- [16] O. Baledent, C. Gondry-Jouet, M.-E. Meyer, G. De Marco, D. Le Gars, M.-C. Henry-Feugeas, and I. Idy-Peretti. “Relationship between cerebrospinal fluid and blood dynamics in healthy volunteers and patients with communicating hydrocephalus”. In: *Investigative Radiology* 39.1 (2004), pp. 45–55.
- [17] B. Bedussi, M. G. Lier, J. W. Bartstra, J. Vos, M. Siebes, E. VanBavel, and E. N.T. P. Bakker. “Clearance from the mouse brain by convection of interstitial fluid towards the ventricular system”. In: *Fluids and Barriers of the CNS* 12.1 (2015), p. 23.
- [18] B. Bedussi, M. Almasian, J. de Vos, E. VanBavel, and E. N.T. P. Bakker. “Paravascular spaces at the brain surface: low resistance pathways for cerebrospinal fluid flow”. In: *Journal of Cerebral Blood Flow & Metabolism* 38.4 (2017), pp. 719–726.
- [19] C. H. L. Beentjes and R. E. Baker. “Quasi-Monte Carlo methods applied to tau-leaping in stochastic biological systems”. In: *Bulletin of Mathematical Biology* (2018), pp. 1–29.
- [20] G. Benfatto, G. Gallavotti, and F. Nicolò. “Elliptic equations and Gaussian processes”. In: *Journal of Functional Analysis* 36.3 (1980), pp. 343–400.
- [21] C. Bernardi. “Optimal finite element interpolation on curved domains”. In: *SIAM Journal on Numerical Analysis* 26.5 (1989), pp. 1212–1240.
- [22] C. Bernardi and V. Girault. “A local regularization operator for triangular and quadrilateral finite elements”. In: *SIAM Journal on Numerical Analysis* 35.5 (1998), pp. 1893–1916.
- [23] G. Beylkin, R. Coifman, and V. Rokhlin. “Wavelets in numerical analysis”. In: *Wavelets and their Applications*. Vol. 181. Jones & Bartlett Boston, 1992, pp. 1–28. ISBN: 978-0-8672-0225-0.
- [24] J. Biehler, M. W. Gee, and W. A. Wall. “Towards efficient uncertainty quantification in complex and large-scale biomechanical problems based on a Bayesian multi-fidelity scheme”. In: *Biomechanics and Modelling in Mechanobiology* 14 (2015), pp. 489–513.
- [25] C. Bierig and A. Chernov. “Approximation of probability density functions by the multilevel Monte Carlo maximum entropy method”. In: *Journal of Computational Physics* 314 (2016), pp. 661–681.

- [26] L. E. Bilston, D. F. Fletcher, A. R. Brodbelt, and M. A. Stoodley. “Arterial pulsation-driven cerebrospinal fluid flow in the perivascular space: a computational model”. In: *Computer Methods in Biomechanics and Biomedical Engineering* 6.4 (2003), pp. 235–241.
- [27] D. Bolin, K. Kirchner, and M. Kovács. “Numerical solution of fractional elliptic stochastic PDEs with spatial white noise”. In: *IMA Journal of Numerical Analysis* dry091 (2018), pp. 1–23.
- [28] D. Bolin, K. Kirchner, and M. Kovács. “Weak convergence of Galerkin approximations for fractional elliptic stochastic PDEs with spatial white noise”. In: *BIT Numerical Mathematics* 58.4 (2018), pp. 881–906.
- [29] D. Bolin and F. Lindgren. “Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping”. In: *Annals of Applied Statistics* 5.1 (2011), pp. 523–550.
- [30] A. Bonito and J. E. Pasciak. “Numerical approximation of fractional powers of elliptic operators”. In: *Mathematics of Computation* 84.295 (2015), pp. 2083–2110.
- [31] S. Brenner and R. L. Scott. *The Mathematical Theory of Finite Element Methods*. Vol. 15. Springer Science and Business Media, 1994. ISBN: 978-1-4757-4340-1.
- [32] R. Buckdahn and E. Pardoux. “Monotonicity methods for white noise driven quasi-linear SPDEs”. In: *Diffusion Processes and Related Problems in Analysis*. Vol. 1. Birkhäuser, 1990, pp. 219–233. ISBN: 978-1-4684-0566-8.
- [33] K. Burrage, N. Hale, and D. Kay. “An efficient implicit FEM scheme for fractional-in-space reaction-diffusion equations”. In: *SIAM Journal on Scientific Computing* 34.4 (2012), pp. 2145–2172.
- [34] L. Caffarelli and L. Silvestre. “An extension problem related to the fractional Laplacian”. In: *Communications in Partial Differential Equations* 32.8 (2007), pp. 1245–1260.
- [35] R. E. Caflish, W. Morokoff, and A. Owen. “Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension”. In: *Journal of Computational Finance* 1 (1997), pp. 27–46.
- [36] Y. Cao, R. Zhang, and K. Zhang. “Finite element and discontinuous Galerkin method for stochastic Helmholtz equation in two-and three-dimensions”. In: *Journal of Computational Mathematics* 26.5 (2008), pp. 702–715.
- [37] R. O. Carare, M. Bernardes-Silva, T. A. Newman, A. M. Page, J. A. R. Nicoll, V. H. Perry, and R. O. Weller. “Solutes, but not cells, drain from the brain parenchyma along basement membranes of capillaries and arteries: significance for cerebral amyloid angiopathy and neuroimmunology”. In: *Neuropathology and Applied Neurobiology* 34.2 (2008), pp. 131–144.
- [38] E. Casas. “ $L^2$  estimates for the finite element method for the Dirichlet problem with singular data”. In: *Numerische Mathematik* 47 (1985), pp. 627–632.

- [39] J. Charrier, R. Scheichl, and A. L. Teckentrup. “Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods”. In: *SIAM Journal of Numerical Analysis* 51.1 (2013), pp. 322–352.
- [40] X. Chen, G. W. Astary, H. Sepulveda, T. H. Mareci, and M. Sarntinoranont. “Quantitative assessment of macromolecular concentration during direct infusion into an agarose hydrogel phantom using contrast-enhanced MRI”. In: *Magnetic Resonance Imaging* 26.10 (2008), pp. 1433–1441.
- [41] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. “Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients”. In: *Computing and Visualization in Science* 14.1 (2011), pp. 3–15.
- [42] S. G. Cox and K. Kirchner. *Regularity and convergence analysis in Sobolev and Hölder spaces for generalized Whittle-Matérn fields*. 2019. arXiv: 1904.06569v1.
- [43] M. Croci, M. B. Giles, M. E. Rognes, and P. E. Farrell. “Efficient white noise sampling and coupling for multilevel Monte Carlo with non-nested meshes”. In: *SIAM/ASA Journal on Uncertainty Quantification* 6.4 (2018), pp. 1630–1655.
- [44] M. Croci, V. Vinje, and M. E. Rognes. “Uncertainty quantification of parenchymal tracer distribution using random diffusion and convective velocity fields”. In: *Fluids and Barriers of the CNS* 16 (32 2019), pp. 1–21.
- [45] I. Daubechies. “Orthonormal bases of compactly supported wavelets”. In: *Communications on Pure and Applied Mathematics* 41.7 (1988), pp. 909–996.
- [46] J. W. Demmel. *Applied Numerical Linear Algebra*. SIAM, 1997. ISBN: 978-0-89871-389-3.
- [47] J. Dick, F. Y. Kuo, and I. H. Sloan. “High-dimensional integration: the quasi-Monte Carlo way”. In: *Acta Numerica* 22 (2013), pp. 133–288.
- [48] A. K. Diem, M. M. G. Sharp, M. Gatherer, N. W. Bressloff, R. O. Carare, and G. Richardson. “Arterial pulsations cannot drive intramural periarterial drainage: significance for A $\beta$  drainage”. In: *Frontiers in Neuroscience* 11.Aug (2017), p. 475.
- [49] C. R. Dietrich and G. N. Newsam. “Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix”. In: *SIAM Journal on Scientific Computing* 18.4 (1997), pp. 1088–1107.
- [50] J. Dölz, H. Harbrecht, and C. Schwab. “Covariance regularity and  $\mathcal{H}$ -matrix approximation for rough random fields”. In: *Numerische Mathematik* 135.4 (2017), pp. 1045–1071.
- [51] D. Drzisga, B. Gmeiner, U. Ude, R. Scheichl, and B. Wohlmuth. “Scheduling massively parallel multigrid for multilevel Monte Carlo methods”. In: *SIAM Journal on Scientific Computing* 39.5 (2017), pp. 873–897.
- [52] Q. Du and T. Zhang. “Numerical approximation of some linear stochastic partial differential equations driven by special additive noises”. In: *SIAM Journal on Numerical Analysis* 40.4 (2002), pp. 1421–1445.

- [53] P. K. Eide and A. Bakken. “The baseline pressure of intracranial pressure (ICP) sensors can be altered by electrostatic discharges”. In: *Biomedical Engineering Online* 10.1 (2011), p. 75.
- [54] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*. 2nd ed. Oxford University Press (UK), 2014. ISBN: 978-0-1996-7879-2.
- [55] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer Science and Business Media, 2013. ISBN: 978-1-4419-1918-2.
- [56] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010. ISBN: 0-8218-4974-3.
- [57] R. D. Falgout and U. M. Yang. “*hypre*: A library of high performance preconditioners”. In: *International Conference on Computational Science*. Springer, 2002, pp. 632–641.
- [58] Q. Fang. “Mesh-based Monte Carlo method using fast ray-tracing in Plücker coordinates”. In: *Biomedical Optics Express* 1.1 (2010), pp. 165–175.
- [59] P. E. Farrell. “Galerkin projection of discrete fields via supermesh construction”. PhD thesis. Imperial College London, 2009.
- [60] P. E. Farrell and J. R. Maddison. “Conservative interpolation between volume meshes by local Galerkin projection”. In: *Computer Methods in Applied Mechanics and Engineering* 200 (2011), pp. 89–100.
- [61] P. E. Farrell, M. D. Piggott, C. C. Pain, G. J. Gorman, and C. R. Wilson. “Conservative interpolation between unstructured meshes via supermesh construction”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (2009), pp. 2632–2642.
- [62] M. Feischl, F. Y. Kuo, and I. H. Sloan. “Fast random field generation with H-matrices”. In: *Numerische Mathematik* 140 (2018), pp. 639–676.
- [63] C. Forbes, M. Evans, N. Hastings, and B. Peacock. *Statistical Distributions*. John Wiley & Sons, 2011. ISBN: 978-0-470-39063-4.
- [64] R. Frič and P. K. Eide. “Comparison of pulsatile and static pressures within the intracranial and lumbar compartments in patients with Chiari malformation type I: a prospective observational study”. In: *Acta Neurochirurgica* 157.8 (2015), pp. 1411–1423.
- [65] C. Geuzaine and J.-F. Remacle. “Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities”. In: *International Journal for Numerical Methods in Engineering* 79.11 (2009), pp. 1309–1331.
- [66] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, 2015. ISBN: 3-540-41160-7.
- [67] M. B. Giles and B. J. Waterhouse. “Multilevel quasi-Monte Carlo path simulation”. In: *Advanced Financial Modelling (Radon Series on Computational and Applied Mathematics)*. 8. De Gruyter, 2009, pp. 165–181. ISBN: 978-3110213133.
- [68] M. B. Giles. “Multilevel Monte Carlo methods”. In: *Acta Numerica* 24 (2015), pp. 259–328.

- [69] M. B. Giles. “Multilevel Monte Carlo path simulation”. In: *Operations Research* 56.3 (2008), pp. 607–617.
- [70] M. B. Giles, T. Nagapetyan, and K. Ritter. “Multilevel Monte Carlo approximation of distribution functions and densities”. In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 267–295.
- [71] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer Science & Business Media, 2013. ISBN: 978-1-4419-1822-2.
- [72] A. Goriely, M. G. D. Geers, G. A. Holzapfel, J. Jayamohan, A. Jérusalem, S. Sivaloganathan, W. Squier, J. A. W. van Dommelen, S. Waters, and E. Kuhl. “Mechanics of the brain: perspectives, challenges, and opportunities”. In: *Biomechanics and Modelling in Mechanobiology* 14.5 (2015), pp. 931–965.
- [73] I. G. Graham, F. Y. Kuo, D Nuyens, R Scheichl, and I. H. Sloan. “Analysis of circulant embedding methods for sampling stationary random fields”. In: *SIAM Journal on Numerical Analysis* 56.3 (2018), pp. 1871–1895.
- [74] I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl, and I. H. Sloan. “Circulant embedding with QMC: analysis for elliptic PDE with lognormal coefficients”. In: *Numerische Mathematik* 140 (2018), pp. 479–511.
- [75] I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, C. Schwab, and I. H. Sloan. “Quasi-Monte Carlo finite element methods for elliptic PDEs with log-normal random coefficients”. In: *Numerische Mathematik* 131.2 (2013), pp. 329–368.
- [76] I. Graham, F. Kuo, D. Nuyens, R. Scheichl, and I. Sloan. “Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications”. In: *Journal of Computational Physics* 230.10 (2011), pp. 3668–3694.
- [77] M. A. Green, L. E. Bilston, and R. Sinkus. “*In vivo* brain viscoelastic properties measured by magnetic resonance elastography”. In: *NMR in Biomedicine* 21.7 (2008), pp. 755–764.
- [78] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. SIAM, 2011. ISBN: 978-1-611972-02-3.
- [79] L. Guo, J. C. Vardakas, T. Lassila, M. Mitolo, N. Ravikumar, D. Chou, M. Lange, A. Sarrami-Foroushani, B. J. Tully, Z. A. Taylor, and Others. “Subject-specific multi-poroelastic model for exploring the risk factors associated with the early stages of Alzheimer’s disease”. In: *Interface Focus* 8.1 (2018), p. 15.
- [80] W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*. Springer-Verlag, 1992. ISBN: 978-3-662-54960-5.
- [81] W. Hackbusch. *Hierarchical Matrices: Algorithms and Analysis*. Springer, 2015, p. 511. ISBN: 978-3-662-47323-8.
- [82] A.-L. Haji-Ali, F. Nobile, E. von Schwerin, and R. Tempone. “Optimization of mesh hierarchies in multilevel Monte Carlo samplers”. In: *Stochastics and Partial Differential Equations Analysis and Computations* 4.1 (2016), pp. 76–112.
- [83] N. Hale, N. J. Higham, and L. N. Trefethen. “Computing  $A^\alpha$ ,  $\log(A)$  and related matrix functions by contour integrals”. In: *SIAM Journal on Numerical Analysis* 46.5 (2008), pp. 2505–2523.

- [84] H. Harbrecht, M. Peters, and R. Schneider. “On the low-rank approximation by the pivoted Cholesky decomposition”. In: *Applied Numerical Mathematics* 62 (2012), pp. 428–440.
- [85] I. F. Harrison, B. Siow, A. B. Akilo, P. G. Evans, O. Ismail, Y. Ohene, P. Navahandi, D. L. Thomas, M. F. Lythgoe, and J. A. Wells. “Non-invasive imaging of CSF-mediated brain clearance pathways via assessment of perivascular fluid movement with DTI MRI”. In: *eLife* 7.e34028 (2018), p. 8.
- [86] P. Hauseux, J. S. Hale, S. Cotin, and S. P. A. Bordas. “Quantifying the uncertainty in a hyperelastic soft tissue model with stochastic parameters”. In: *Applied Mathematical Modelling* 62 (2018), pp. 86–102.
- [87] S. Heinrich. “Multilevel Monte Carlo Methods”. In: *Large-Scale Scientific Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 58–67.
- [88] L. Herrmann and C. Schwab. “Multilevel quasi-Monte Carlo integration with product weights for elliptic PDEs with lognormal coefficients”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 53.5 (2019), pp. 1507–1552.
- [89] T. Hida, H.-H. Kuo, J. Potthoff, and L. Streit. *White Noise: an Infinite Dimensional Calculus*. 1st ed. Springer Science & Business Media, 1993. ISBN: 978-90-481-4260-6.
- [90] S. B. Hladky and M. A. Barrand. “Mechanisms of fluid movement into, through and out of the brain: evaluation of the evidence”. In: *Fluids and Barriers of the CNS* 11.1 (2014), p. 26.
- [91] K. E. Holter, M. Kuchta, and K.-A. Mardal. “Sub-voxel perfusion modeling in terms of coupled 3d-1d problem”. In: (2018). arXiv: 1803.04896.
- [92] K. E. Holter, B. Kehlet, A. Devor, T. J. Sejnowski, A. M. Dale, S. W. Omholt, O. P. Ottersen, E. A. Nagelhus, K.-A. Mardal, and K. H. Pettersen. “Interstitial solute transport in 3D reconstructed neuropil occurs by diffusion rather than bulk flow”. In: *Proceedings of the National Academy of Sciences* 114.37 (2017), pp. 9894–9899.
- [93] J. J. Iliff, M. Wang, Y. Liao, B. A. Plogg, W. Peng, G. A. Gundersen, and Others. “A paravascular pathway facilitates CSF flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid beta”. In: *Science Translational Medicine* 4.147 (2012), 147ra111.
- [94] J. J. Iliff, M. Wang, D. M. Zeppenfeld, A. Venkataraman, B. A. Plog, Y. Liao, R. Deane, and M. Nedergaard. “Cerebral arterial pulsation drives paravascular CSF–interstitial fluid exchange in the murine brain”. In: *Journal of Neuroscience* 33.46 (2013), pp. 18190–18199.
- [95] H. Inaba and B. D. Tapley. “Generalized random processes: a theory and the white Gaussian process”. In: *SIAM Journal on Control* 13.3 (1975), pp. 719–735.
- [96] N. A. Jessen, A. S. F. Munk, I. Lundgaard, and M. Nedergaard. “The glymphatic system: a beginner’s guide”. In: *Neurochemical Research* 40.12 (2015), pp. 2583–2599.

- [97] B.-J. Jin, A. J. Smith, and A. S. Verkman. “Spatial model of convective solute transport in brain extracellular space does not support a ‘glymphatic’ mechanism”. In: *Journal of General Physiology* 148.6 (2016), pp. 489–501.
- [98] S. Joe and F. Y. Kuo. “Constructing Sobol sequences with better two-dimensional projections”. In: *SIAM Journal on Scientific Computing* 30.5 (2008), pp. 2635–2654.
- [99] H. Jung. “Über den kleinsten Kreis, der eine ebene Figur einschließt.” In: *Journal Für die Reine und Angewandte Mathematik* 137 (1910), pp. 310–313.
- [100] H. Jung. “Über die kleinste Kugel, die eine räumliche Figur einschließt”. In: *Journal Für die Reine und Angewandte Mathematik* 123 (1901), pp. 241–257.
- [101] K. G. Kapoor, S. E. Katz, D. M. Grzybowski, and M. Lubow. “Cerebrospinal fluid outflow: an evolving perspective”. In: *Brain Research Bulletin* 77.6 (2008), pp. 327–334.
- [102] B. N. Khoromsjij, A Litvinenko, H. G. Matthies, A Litvinenko, and H. G. Matthies. “Application of hierarchical matrices for computing the Karhunen-Loève expansion”. In: *Computing* 84.1-2 (2009), pp. 49–67.
- [103] U Khristenko, L Scarabosio, P Swierczynski, E Ullmann, and B Wohlmuth. *Analysis of boundary effects on PDE-based sampling of Whittle-Matérn random fields*. 2018. arXiv: 1809.07570v1.
- [104] V. Kiviniemi, X. Wang, V. Korhonen, T. Keinänen, T. Tuovinen, J. Autio, P. LeVan, S. Keilholz, Y.-F. Zang, J. Hennig, and Others. “Ultra-fast magnetic resonance encephalography of physiological brain activity—Glymphatic pulsation mechanisms?” In: *Journal of Cerebral Blood Flow & Metabolism* 36.6 (2016), pp. 1033–1045.
- [105] I. Kiyosi. “Stationary random distributions”. In: *Memoirs of the College of Science, University of Kyoto. Series A: Mathematics* 28.3 (1954), pp. 209–223.
- [106] J. Krejza, M. Arkuszewski, S. E. Kasner, J. Weigle, A. Ustymowicz, R. W. Hurst, B. L. Cucchiara, and S. R. Messe. “Carotid artery diameter in men and women and the relation to body and neck size”. In: *Stroke* 37.4 (2006), pp. 1103–1105.
- [107] S. Krumscheid and F. Nobile. “Multilevel Monte Carlo approximation of functions”. In: *SIAM/ASA Journal on Uncertainty Quantification* 6.3 (2018), pp. 1256–1293.
- [108] F. Y. Kuo, C. Schwab, and I. H. Sloan. “Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients”. In: *Foundations of Computational Mathematics* 15.2 (2015), pp. 411–449.
- [109] F. Y. Kuo, R. Scheichl, C. Schwab, I. H. Sloan, and E. Ullmann. “Multilevel quasi-Monte Carlo methods for lognormal diffusion problems”. In: *Mathematics of Computation* 86.308 (2017), pp. 2827–2860.
- [110] F. Y. Kuo and D. Nuyens. *A practical guide to quasi-Monte Carlo methods*. Tech. rep. KU Leuven, 2016, pp. 1–50.

- [111] P. Lacolley, P. Challande, S. Boumaza, G. Cohuet, S. Laurent, P. Boutouyrie, J.-A. Grimaud, D. Paulin, J.-M. D. Lamaziere, and Z. Li. “Mechanical properties and structure of carotid arteries in mice lacking desmin”. In: *Cardiovascular Research* 51.1 (2001), pp. 178–187.
- [112] P. L’Ecuyer and C. Lemieux. “Variance reduction via lattice rules”. In: *Management Science* 46.9 (2000), pp. 1214–1235.
- [113] V. Lemaire and G. Pagès. “Multilevel Richardson-Romberg extrapolation”. In: *Bernoulli* 23.4A (2017), pp. 2643–2692.
- [114] C. Lemieux. *Monte Carlo and Quasi-Monte Carlo Sampling*. 1st ed. Springer Series in Statistics. Springer-Verlag New York, 2009. ISBN: 978-0-387-78164-8.
- [115] X. Li and J. Liu. *A multilevel approach towards unbiased sampling of random elliptic partial differential equations*. 2016. arXiv: 1605.06349.
- [116] F. Lindgren, H. Rue, and J. Lindström. “An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.4 (2009), pp. 423–498.
- [117] A. A. Linninger, K. Tangen, C.-Y. Hsu, and D. Frim. “Cerebrospinal fluid mechanics and its coupling to cerebrovascular dynamics”. In: *Annual Review of Fluid Mechanics* 48.1 (2016), pp. 219–257.
- [118] A. A. Linninger, M. R. Somayaji, M. Mekarski, and L. Zhang. “Prediction of convection-enhanced drug delivery to the human brain”. In: *Journal of Theoretical Biology* 250.1 (2008), pp. 125–138.
- [119] A. Litvinenko, Y. Sun, M. G. Genton, and D. E. Keyes. “Likelihood approximation with hierarchical matrices for large spatial datasets”. In: *Computational Statistics & Data Analysis* 137 (2019), pp. 115–132.
- [120] A. Logg, K.-A. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Element Method: the FEniCS Book*. Vol. 84. Springer Science & Business Media, 2012. ISBN: 978-3-642-23098-1.
- [121] J. R. Maddison, P. E. Farrell, and I. Panourgias. *Parallel supermeshing for multimesh modelling*. Tech. rep. Archer, eCSE03-08, 2013.
- [122] H. Mestre, J. Tithof, T. Du, W. Song, W. Peng, A. M. Sweeney, G. Olveda, J. H. Thomas, M. Nedergaard, and D. H. Kelley. “Flow of cerebrospinal fluid is driven by arterial pulsations and is reduced in hypertension”. In: *Nature Communications* 9.1 (2018), p. 4878.
- [123] W. J. Morokoff and R. E. Caflisch. “Quasi-Monte Carlo integration”. In: *Journal of Computational Physics* 122.2 (1995), pp. 218–230.
- [124] M. P. Murphy and H. LeVine III. “Alzheimer’s disease and the amyloid- $\beta$  peptide”. In: *Journal of Alzheimer’s Disease* 19.1 (2010), pp. 311–323.
- [125] R. B. Nelsen. *An Introduction to Copulas*. Springer Science & Business Media, 2007. ISBN: 978-0-387-28659-4.
- [126] J. Nečas. *Les Méthodes Directes en Théorie des Équations Elliptiques*. 1st ed. Masson, Academia, 1967. ISBN: 978-2-2255-4868-0.

- [127] C. Nicholson. “Diffusion and related transport mechanisms in brain tissue”. In: *Reports on Progress in Physics* 64.7 (2001), p. 815.
- [128] C. Nilsson, F. Stahlberg, C. Thomsen, O. Henriksen, M. Herning, and C. Owman. “Circadian variation in human cerebrospinal fluid production measured by magnetic resonance imaging”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 262.1 (1992), pp. 20–24.
- [129] J. A. Nitsche and A. H. Schatz. “Interior estimates for Ritz-Galerkin methods”. In: *Mathematics of Computation* 28.128 (1974), pp. 937–958.
- [130] D. Orešković, M. Radoš, and M. Klarica. “New concepts of cerebrospinal fluid physiology and development of hydrocephalus”. In: *Pediatric Neurosurgery* 52.6 (2017), pp. 417–425.
- [131] D. Orešković, M. Radoš, and M. Klarica. “Role of choroid plexus in cerebrospinal fluid hydrodynamics”. In: *Neuroscience* 354 (2017), pp. 69–87.
- [132] S. Osborn, P. S. Vassilevski, and U. Villa. *A multilevel, hierarchical sampling technique for spatially correlated random fields*. 2017. arXiv: 1703.08498.
- [133] S. Osborn, P. Zulian, T. Benson, U. Villa, R. Krause, and P. S. Vassilevski. *Scalable hierarchical PDE sampler for generating spatially correlated random fields using non-matching meshes*. 2017. arXiv: 1712.06758.
- [134] A. B. Owen. “Randomly permuted  $(t, m, s)$ -nets and  $(t, s)$ -sequences.” In: *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*. Springer, New York, NY, 1995, pp. 299–317. ISBN: 978-0-387-94577-4.
- [135] A. B. Owen. “Scrambled net variance for integrals of smooth functions”. In: *Annals of Statistics* 25.4 (1997), pp. 1541–1562.
- [136] A. B. Owen. “Scrambling Sobol’ and Niederreiter-Xing points”. In: *Journal of Complexity* 14 (1998), pp. 466–489.
- [137] A. B. Owen. “Variance and discrepancy with alternative scramblings”. In: *ACM Transactions on Modeling and Computer Simulation* 13.4 (2003), pp. 363–378.
- [138] P. Whittle. “On stationary processes in the plane”. In: *Biometrika* 41.3-4 (1954), pp. 434–449.
- [139] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and Others. “Scikit-learn: Machine learning in Python”. In: *Journal of Machine Learning Research* 12.Oct (2011), pp. 2825–2830.
- [140] S. Pilipović and D. Seleši. “Structure theorems for generalized random processes”. In: *Acta Mathematica Hungarica* 117.3 (2007), pp. 251–274.
- [141] R. Potsepaev and C. L. Farmer. “Application of stochastic partial differential equations to reservoir property modelling”. In: *ECMOR XII-12th European Conference on the Mathematics of Oil Recovery*. Vol. 2. September. 2014. ISBN: 978-1-6343-9168-9.
- [142] A. Quaglino, S. Pezzuto, and R. Krause. *Generalized multifidelity Monte Carlo estimators*. 2018. arXiv: 1807.10521v1.

- [143] Z. Rajna, L. Raitamaa, T. Tuovinen, J. Heikkilä, V. Kiviniemi, and T. Seppänen. “3D multi-resolution optical flow analysis of cardiovascular pulse propagation in human brain”. In: *IEEE Transactions on Medical Imaging* 38.9 (2019), p. 9.
- [144] L. Ray, J. J. Iliff, and J. J. Heys. “Analysis of convective and diffusive transport in the brain interstitium”. In: *Fluids and Barriers of the CNS* 16.1 (2019), p. 6.
- [145] M. L. Rennels, T. F. Gregory, O. R. Blaumanis, K. Fujimoto, and P. A. Grady. “Evidence for a ‘paravascular’ fluid circulation in the mammalian central nervous system, provided by the rapid distribution of tracer protein throughout the brain from the subarachnoid space”. In: *Brain Research* 326.1 (1985), pp. 47–63.
- [146] G. Ringstad, S. A. S. Vatnehol, and P. K. Eide. “Glymphatic MRI in idiopathic normal pressure hydrocephalus”. In: *Brain* 140.10 (2017), pp. 2691–2705.
- [147] G. Ringstad, L. M. Valnes, A. M. Dale, A. H. Pripp, S.-A. S. Vatnehol, K. E. Emblem, K.-A. Mardal, and P. K. Eide. “Brain-wide glymphatic enhancement and clearance in humans assessed with MRI”. In: *JCI Insight* 3.13 (2018), p. 16.
- [148] G. Roth and U. Dicke. “Evolution of the brain and intelligence”. In: *Trends in Cognitive Sciences* 9.5 (2005), pp. 250–257.
- [149] R. Scheichl. *Personal communication*. 2017.
- [150] M. Scheuerer. “Regularity of the sample paths of a general second order random field”. In: *Stochastic Processes and their Applications* 120 (2010), pp. 1879–1897.
- [151] C. Schwab and R. A. Todor. “Karhunen-Loéve approximation of random fields by generalized fast multipole methods”. In: *Journal of Computational Physics* 217.1 (2006), pp. 100–122.
- [152] C. Schwab. *p- and hp- Finite Element Methods. Theory and Applications in Solid and Fluid Mechanics*. Clarendon Press, 1998. ISBN: 978-0-1985-0390-3.
- [153] K. Shahim, J.-M. Drezet, B. A. Martin, and S. Momjian. “Ventricle equilibrium position in healthy and normal pressure hydrocephalus brains using an analytical model”. In: *Journal of Biomechanical Engineering* 134.4 (2012), p. 10.
- [154] K. Shahim, J.-M. Drezet, J.-F. Molinari, R. Sinkus, and S. Momjian. “Finite element analysis of normal pressure hydrocephalus: influence of CSF content and anisotropy in permeability”. In: *Applied Bionics and Biomechanics* 7.3 (2010), pp. 187–197.
- [155] M. K. Sharp, R. O. Carare, and B. A. Martin. “Dispersion in porous media in oscillatory flow between flat plates: applications to intrathecal, periarterial and paraarterial solute transport in the central nervous system”. In: *Fluids and Barriers of the CNS* 16.1 (2019), p. 13.
- [156] G. D. Silverberg, M. Mayo, T. Saul, E. Rubenstein, and D. McGuire. “Alzheimer’s disease, normal pressure hydrocephalus, and senescent changes in CSF circulatory physiology: a hypothesis”. In: *The Lancet Neurology* 2.8 (2003), pp. 506–511.

- [157] D Simpson, J. B. Illian, F Lindgren, S. H. Sørbye, H Rue, and J Illian. “Going off grid: Computationally efficient inference for log-Gaussian Cox processes”. In: *Biometrika* 20 (2015), pp. 1–22.
- [158] A. J. Smith and A. S. Verkman. “The ‘glymphatic’ mechanism for solute clearance in Alzheimer’s disease: game changer or unproven speculation?” In: *The FASEB Journal* 32.2 (2017), pp. 543–551.
- [159] A. J. Smith, X. Yao, J. A. Dix, B.-J. Jin, and A. S. Verkman. “Test of the ‘glymphatic’ hypothesis demonstrates diffusive and aquaporin-4-independent solute transport in rodent brain parenchyma”. In: *eLife* 6.e27679 (2017), p. 16.
- [160] I. M. Sobol’, D. Asotsky, A. Kreinin, and S. Kucherenko. “Construction and comparison of high-dimensional Sobol’ generators”. In: *Wilmott* Nov. 2011.56 (2011), pp. 64–79.
- [161] K. H. Støverud, M. Darcis, R. Helmig, and S. M. Hassanzadeh. “Modeling concentration distribution and deformation during convection-enhanced drug delivery into brain tissue”. In: *Transport in Porous Media* 92.1 (2012), pp. 119–143.
- [162] G. Strang. *Introduction to Linear Algebra*. 5th ed. Wellesley-Cambridge Press, 2016. ISBN: 978-09802327-7-6.
- [163] T. J. Sullivan. *Introduction to Uncertainty Quantification*. Vol. 63. Springer, 2015. ISBN: 978-3-319-23394-9.
- [164] E. Syková and C. Nicholson. “Diffusion in brain extracellular space”. In: *Physiological Reviews* 88.4 (2008), pp. 1277–1340.
- [165] J. M. Tarasoff-Conway, R. O. Carare, R. S. Osorio, L. Glodzik, T. Butler, E. Fieremans, L. Axel, H. Rusinek, C. Nicholson, B. V. Zlokovic, and Others. “Clearance systems in the brain - implications for Alzheimer disease”. In: *Nature Reviews Neurology* 11.8 (2015), pp. 457–470.
- [166] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann. “Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients”. In: *Numerische Mathematik* 125.3 (2013), pp. 569–600.
- [167] V. Thomée. “On positivity preservation in some finite element methods for the heat equation”. In: *International Conference on Numerical Methods and Applications*. Springer. 2014, pp. 13–24.
- [168] G. F. Tóth and W. Kuperberg. “A survey of recent results in the theory of packing and covering”. In: *New Trends in Discrete and Computational Geometry*. Berlin, Heidelberg: Springer, 1993, pp. 251–279.
- [169] U. Trottenberg, C. W. Oosterlee, and A. Schuller. *Multigrid*. Academic Press, 2000. ISBN: 0-12-701070-X.
- [170] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen. “High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity”. In: *Magnetic Resonance in Medicine* 48.4 (2002), pp. 577–582.
- [171] B. J. Tully and Y. Ventikos. “Cerebral water transport using multiple-network poroelastic theory: application to normal pressure hydrocephalus”. In: *Journal of Fluid Mechanics* 667 (2011), pp. 188–215.

- [172] G. F. Vindedal, A. E. Thoren, V. Jensen, A. Klungland, Y. Zhang, M. J. Holtzman, O. P. Ottersen, and E. A. Nagelhus. “Removal of aquaporin-4 from glial and ependymal membranes causes brain water accumulation”. In: *Molecular and Cellular Neuroscience* 77 (2016), pp. 47–52.
- [173] A. J. Wathen. “Realistic eigenvalue bounds for the Galerkin mass matrix”. In: *IMA Journal of Numerical Analysis* 7.August (1987), pp. 449–457.
- [174] R. O. Weller. “Pathology of cerebrospinal fluid and interstitial fluid of the CNS: significance for Alzheimer’s disease, prion disorders and multiple sclerosis”. In: *Journal of Neuropathology & Experimental Neurology* 57.10 (1998), pp. 885–894.
- [175] J. G. Wendel. “Note on the gamma function”. In: *The American Mathematical Monthly* 55.9 (1948), pp. 563–564.
- [176] A. T. A. Wood and G. Chan. “Simulation of stationary Gaussian processes in  $[0, 1]^d$ ”. In: *Journal of Computational and Graphical Statistics* 3.4 (1994), pp. 409–432.
- [177] J. H. Wood. *Neurobiology of Cerebrospinal Fluid*. 2nd ed. Springer Science & Business Media, 2013. ISBN: 978-1-4615-9271-6.
- [178] World Health Organization. *Dementia: key facts*. 2016. URL: [www.who.int/mediacentre/factsheets/fs362/en/](http://www.who.int/mediacentre/factsheets/fs362/en/).
- [179] L. Xie, H. Kang, Q. Xu, M. J. Chen, Y. Liao, M. Thiagarajan, J. O’Donnell, D. J. Christensen, C. Nicholson, J. J. Iliff, and Others. “Sleep drives metabolite clearance from the adult brain”. In: *Science* 342.6156 (2013), pp. 373–377.
- [180] Z. Ye, Y. Liu, X. Wang, X. Chen, C. Lin, Y. Su, T. Wang, Y. Tang, Y. Wu, and C. Qin. “A rhesus monkey model of common carotid stenosis”. In: *International Journal of Clinical and Experimental Medicine* 9.9 (2016), pp. 17487–17497.
- [181] H. Zetterberg, R. Lautner, T. Skillback, C. Rosen, P. Shahim, N. Mattsson, and K. Blennow. “CSF in Alzheimer’s disease”. In: *Advances in Clinical Chemistry* 65 (2014), pp. 143–172.
- [182] Z. Zhang, B. Rozovskii, and G. E. Karniadakis. “Strong and weak convergence order of finite element methods for stochastic PDEs with spatial white noise”. In: *Numerische Mathematik* 134 (2016), pp. 61–89.