Multilevel Solvers for Stochastic Fluid Flows

Kumar, Prashant

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Multilevel Solvers

for Stochastic Fluid Flows

Multilevel Solvers

for Stochastic Fluid Flows

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof. dr. ir. T. H. J. J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Tuesday 16 July 2019 at 10:00 o'clock

by

Prashant KUMAR

Master of Science in Exploration Geophysics,
Indian Institute of Technology Kharagpur, India,
born in Ranchi, India.

This dissertation has been approved by the promotors

Promotor: Dr. R. P. Dwight
Promotor: Prof. dr. ir. C. W. Oosterlee

Composition of the doctoral committee:

Rector Magnificus	chairperson
Dr. R. P. Dwight	Delft University of Technology, promotor
Prof. dr. ir. C. W. Oosterlee	Delft University of Technology, promotor

Independent members:

Prof. dr. H. Xiao	Virginia Tech, USA
Prof. dr. D. Nuyens	Catholic University of Leuven, Belgium
Prof. dr. S. Hickel	Delft University of Technology
Prof. dr. ir. C. Vuik	Delft University of Technology

Other member:

Prof. dr. F. J. Gaspar	University of Zaragoza, Spain

Dissertation at Delft University of Technology

An electronic version of this dissertation is available at

`http://repository.tudelft.nl/`.

Summary

Uncertainty is ubiquitous in many areas of science and engineering. It may result from the inadequacy of mathematical models to represent the reality or from unknown physical parameters that are required as inputs for these models. Uncertainty may also arise due to the inherent randomness of the system being analyzed. For many problems of practical interest, uncertainty quantification (UQ) can involve computations that are intractable even for the modern supercomputers, if conventional mathematical techniques are utilized. The reason is typically a product of complexity factors associated with many samples needed to compute the statistics, and for each sample, complexity associated with the spatio-temporal scales characteristics to the system.

The main objective of this research is to obtain *multilevel solvers* for stochastic fluid flow problems with high-dimensional uncertainties. In our approach, the complexity arising due to sampling is overcome by the multilevel Monte Carlo (MLMC) method and complexity due to spatio-temporal scales is eliminated via the multigrid solver. Historically, Monte Carlo (MC) type methods have been proven to be the methods of choice for problems with a large uncertainty dimension as they do not suffer from the *curse of dimensionality*. A well-known computational bottleneck associated with the plain MC method is the slow convergence of the sampling error. For problems involving a wide range of space and time scales, ensuring a low mean square error will require a large number of MC samples on a very fine computational mesh making the estimator very expensive. Inspired by the multigrid ideas, the MLMC method generalizes the standard MC to multiple grids, exhibiting an exceptional improvement. The efficiency of the MLMC method comes from solving the problem of interest on a coarse grid and subsequently adding corrections based on finer mesh resolutions. On the coarsest grid, a large number of samples can be computed inexpensively. The corrections computed on finer grids, have smaller variances and can be estimated accurately using only fewer samples. The estimates at different levels are then combined using a telescopic sum.

A fast and robust black-box solver is very important for designing an efficient MLMC estimator. Therefore, combining the MLMC method with a multigrid method seems natural, where the multigrid method is employed for the numerical solution of a partial differential equation (PDE) with uncertainty on a certain scale. Developing a multigrid method for problems with random inputs is challenging, as depending on the random inputs, the multigrid convergence rate may have a high variance or even diverge for certain samples. In this thesis, we also develop multigrid methods that are well suited to an MLMC setting. We pay special attention to the choice of the multigrid components and provide modifications needed in a deterministic multigrid solver to deal with the stochasticity.

We study four different fluid flow problems: single-phase flow in porous media with highly heterogeneous diffusion coefficients; a multi-physics problem involving advection-dominated transport in a coupled Darcy-Stokes system; a nonlinear multi-phase flow in variably saturated porous media; and turbulent flows with high Reynolds number. For all these problems, we encounter uncertainties that are extremely high-dimensional where the unknown physical parameters are modeled as an infinite-dimensional random field or even as an infinite-dimensional random tensor field.

To gain an insight into the MLMC method, we also study the theory and sampling strategies, in Chapter 2. Then, we explore the possibilities for improving the complexity of the MLMC method

using higher-order discretization schemes in Chapter 3. Here, we also describe a framework to combine a full multigrid (FMG) solver with the MLMC estimator using the same hierarchy of grids. This coupling of a FMG solver with the MLMC estimator is highly appealing, as we can obtain linear-scaling UQ methods for many problems.

In Chapter 4, we study a non-standard Local Fourier Analysis (LFA) to predict the asymptotic convergence factors of multigrid solvers for PDEs with random or jumping coefficients. The proposed LFA technique can be used to compute the convergence statistics of multigrid solvers in advance and can be used to optimize the MLMC simulation.

Devising an efficient MLMC method for coupled multi-physics problems is highly challenging. A careful selection of numerical schemes for each subproblem is needed to avoid a high cost per sample without any gain in accuracy. For such systems, the approximation of each subproblem may give rise to different convergence rates depending on the regularity of the solution variables and on the discretization order of the numerical schemes. These aspects are studied in Chapter 5 where we discuss a multigrid MLMC approach for stochastic transport in a coupled Darcy-Stokes system. With the combination of numerical schemes described in this chapter, we are able to obtain an asymptotically optimal MLMC estimator that has the same computational complexity as its deterministic counterpart.

The standard MLMC method relies on a sequence of coarse grids for cost reduction. For stiff nonlinear systems, a convergent solution can only be obtained on relatively finer grids, thereby limiting the full potential of the MLMC method. To deal with this shortcoming, we explore a *parametric continuation* strategy for nonlinear PDEs in Chapter 6. The key idea is to simplify the nonlinearity as we move to coarse grids. This way we circumvent the problem of divergence as simpler nonlinear problems are now solved on coarser levels. To test the effectiveness of this approach, we consider flow in variably saturated porous media that is modeled using the nonlinear Richards' equation. For a widely used solution method based on the modified Picard iteration, we show that the parametric continuation approach can accelerate the convergence of the standard MLMC method for nonlinear problems.

In Chapter 7, we extend the MLMC framework for quantifying *model-form* uncertainties associated with the Reynolds-Averaged Navier-Stokes (RANS) simulations. The RANS equations combined with turbulence closure models are widely utilized in engineering to predict flows at high Reynolds numbers. The turbulence models are typically derived using a combination of physical insight and empirical data-fitting. This modeling approach, although computationally inexpensive introduces large modeling errors in the RANS prediction. This work describes two stochastic extensions for RANS models based on *random eddy viscosity* and a *random Reynolds Stress Tensor* (RST) which are solved using the MLMC method. The random eddy viscosity is obtained by perturbing the baseline eddy viscosity using Gaussian random fields with some prescribed spatial covariance. Similarly, the random RST is derived by perturbing the deterministic baseline RST. For the proposed method, we show that the asymptotic cost of the MLMC estimator does not deteriorate with an increase in uncertainty dimension. Here, we also demonstrate that the two stochastic models are sufficiently general and can reliably bound the possible flow behavior.

This PhD thesis reports on the multilevel approach for UQ of fluid flow problems. Our focus is on the development of fast uncertainty propagation algorithms. We also show the generality of the considered stochastic models for different classes of flow. As a next step, based on the machinery developed in this work, one can further develop efficient data-driven inverse UQ algorithms for parameter calibration and quantification of model errors.

Samenvatting

Onzekerheid is een alom vertegenwoordigd onderdeel van veel technisch-wetenschappelijke vakgebieden en modellen. Voorbeelden van bronnen van onzekerheid zijn ten eerste de fout van een model ten opzichte van de fysische realiteit en ten tweede onbekende parameters waarvan een model afhangt. Het is ook mogelijk dat het fysische model zelf onzekerheid bevat. Het modelleren van onzekerheid staat bekend als onzekerheidskwantificering (afgekort UQ, naar het Engels *Uncertainty Quantification*), maar met conventionele technieken uit dit vakgebied kunnen veel relevante problemen niet worden opgelost. Dit komt doordat de benodigde berekeningen zelfs voor moderne supercomputers te complex zijn. De reden hiervan is tweeledig. Ten eerste kan de onzekerheid alleen gekwantificeerd worden door het onderliggende deterministische model vaak op te lossen en ten tweede is het onderliggende model complex door de vele schalen van het probleem (zowel in ruimte als in tijd).

Het hoofddoel van dit onderzoek is het afleiden van numerieke *multilevel oplossingsmethoden* welke toegepast kunnen worden op stochastische vloeistof dynamica problemen met hoog dimensionale onzekerheden. Onze aanpak bestaat uit het combineren van de multilevel Monte Carlo (MLMC) methode met de multigrid methode, waardoor zowel de problematiek van de vele model evaluaties alsmede de complexiteit van de vele schalen in het model zelf worden tegen gegaan. Het is algemeen bekend dat (gewone) Monte Carlo (MC) methoden effectief omgaan met hoog dimensionale onzekerheden, omdat de hoeveelheid benodigde numerieke oplossingen van het model niet afhangt van het aantal onzekerheden (dit fenomeen staat bekend als de *curse of dimensionality*). Echter ondanks deze kenmerkende eigenschap zijn er veel numerieke oplossingen van het model nodig om tot een accurate schatting te komen. Dit speelt vooral wanneer er vele schalen in het probleem zijn. MLMC gaat dit effect tegen door MC toe te passen op meerdere roosters van het onderliggende model, waardoor er substantieel minder tijd nodig is om tot een oplossing te komen. Het idee is om het onderliggende model vaak numeriek op lossen gebruik makende van een grof rooster en vervolgens de fout die hierdoor gemaakt wordt te corrigeren gebruik makende van oplossingen bepaald met steeds fijner wordende roosters. Het grove rooster kan gebruikt worden om het model vaak numeriek op te lossen, aangezien dit snel en effectief kan. De variantie tussen een oplossing op het grove rooster en een fijn rooster is klein, dus er zijn maar weinig numerieke oplossingen nodig op een fijn rooster om accurate statistieken te berekenen. De correcties worden vervolgens gecombineerd door middel van een telescopische som.

MLMC is een grote verbetering ten opzichte van MC, maar vereist nog steeds een snelle en robuuste numerieke oplossingsmethode voor het onderliggende model. Een voor de hand liggende idee is om de MLMC methode te combineren met een multigrid methode, waarbij de multigrid methode wordt gebruikt voor het oplossen van een partiële differentiaalvergelijking (PDV) op een vooraf bepaalde schaal. Het is niet evident om een multigrid methode te gebruiken wanneer de parameters van de differentiaalvergelijking stochasten zijn, aangezien afhankelijk van de parameters de multigrid methode langzaam (met hoge variantie) of zelfs niet convergeert. In deze dissertatie worden multigrid methodieken ontwikkeld die zeer effectief zijn in combinatie met MLMC. We letten vooral op de juiste keuze van alle onderdelen in multigrid en ontwikkelen aanpassingen daar waar nodig.

We bestuderen vier verschillende vloeistof dynamica problemen: eenfasige stroming in poreuze

media met sterk variabele diffusie; een combinatie van een door advectie gedomineerde stroming en een Darcy–Stokes systeem; een niet-lineaire meerfasige stroming in verzadigde poreuze media; en turbulente stroming met hoog Reynoldsgetal. In alle problemen komen extreem hoog dimensionale onzekerheden voor waarbij de onzekere fysische parameters gemodelleerd worden als oneindig dimensionale stochastische velden of zelfs als oneindig dimensionale stochastische tensorvelden.

Om inzicht te verwerven in MLMC bestuderen we in Hoofdstuk 2 de wiskundige theorie en strategieën om steekproeven te bepalen. Daarna verkennen we in Hoofdstuk 3 de mogelijkheden om de MLMC methode te versnellen gebruik makende van hogere orde discretisatie methodieken. We beschrijven hier ook een raamwerk om *full multigrid* (FMG) te combineren met MLMC, waarbij beide methoden gebruik maken van dezelfde roosterhiërarchie. De koppeling tussen FMG en MLMC is veelbelovend, aangezien de tijdscomplexiteit van de verkregen UQ methodieken in veel gevallen lineair schaalt.

In Hoofdstuk 4 leiden we een nieuwe lokale Fourier analyse (LFA) af om asymptotische convergentie van multigrid oplossingen van PDV's met willekeurige coëfficiënten te voorspellen. De afgeleide LFA techniek kan gebruikt worden om convergentie van multigrid technieken vooraf de kwantificeren en kan gebruikt worden om MLMC simulaties te optimaliseren.

Het is niet eenvoudig een efficiënte MLMC methode te ontwikkelen voor problemen waarin verschillende fysische fenomenen gekoppeld zijn. Het is van belang om nauwkeurig de numerieke schema's te kiezen welke gebruikt worden voor ieder fenomeen of submodel zodat de precisie van ieder onderdeel hetzelfde is. De convergentiesnelheid kan verschillen tussen de onderdelen, afhankelijk van de regulariteit van de oplossing en de discretisatie van het fysische submodel. Deze aspecten worden beschouwd in Hoofdstuk 5, waarin een multigrid MLMC methode wordt afgeleid voor een stochastische stroming in een gekoppeld Darcy–Stokes systeem. De numerieke schema's die worden afgeleid in dit hoofdstuk worden gecombineerd, om zo een asymptotisch optimale MLMC schatter te verkrijgen, welke dezelfde complexiteit heeft als het deterministische probleem.

The standaard MLMC methode maakt gebruik van een serie grove rooster om de effectiviteit te waarborgen. Echter voor stijve niet-lineaire systemen is het alleen mogelijk om een geconvergeerde numerieke oplossing te verkrijgen op relatief fijne roosters, waardoor de effectiviteit van MLMC afneemt. Om dit probleem op te lossen verkennen we in Hoofdstuk 6 *parametrische voortzetting*, wat toegepast kan worden op niet-lineaire PDV's. Het idee is om de niet-lineariteit voor grove roosters te vereenvoudigen. Hiermee kan sneller een geconvergeerde numerieke oplossing verkregen, aangezien het probleem aanzienlijk versimpeld is. Om de effectiviteit van deze aanpak te testen wordt deze toepast op stroming in een verzadigd poreus medium, wat gemodelleerd wordt door middel van de niet-lineaire vergelijking van Richards. We gebruiken een methode welke gebaseerd is op de veelgebruikte Picard iteratie om aan te tonen dat parametrische voortzetting de convergentie van MLMC in niet-lineaire problemen aanzienlijk kan versnellen.

In Hoofdstuk 7 bereiden we het MLMC raamwerk uit om de *modelonzekerheden* van de stationaire Navier–Stokes vergelijkingen voor de gemiddelde stroming te kwantificeren. Deze vergelijkingen staat bekend als de de *Reyolds-averaged Navier–Stokes* vergelijkingen, oftewel de RANS vergelijkingen. De RANS vergelijkingen worden vaak in combinatie met een turbulentie-model gebruikt om het gedrag van vloeistoffen met een hoog Reynoldsgetal te voorspellen. De turbulentie-modellen worden normaal gesproken verkregen door fysisch inzicht en empirische data-analyses. Deze aanpak zorgt voor modellen die snel numeriek opgelost kunnen worden, maar een grote inherente fout bevatten. In Hoofdstuk 7 worden twee stochastische expansies voor RANS modellen beschreven, welke gebaseerd zijn op een *willekeurige eddy viscositeit* en een *willekeurige Reynolds stress tensor* (RST). Deze worden opgelost gebruik makende van de MLMC methode. De willekeurige eddy viscositeit is verkregen door de gebruikelijke waarde van de eddy viscositeit op te tellen bij een hoog-dimensionaal Gaussisch proces. Op dezelfde wijze wordt de willekeurige RST verkregen. We laten zien dat de benodigde tijd om de nieuwe MLMC methodieken toe te passen asymptotisch

niet toeneemt wanneer de hoeveelheid onzekere parameters toeneemt. We laten ook zien dat de twee stochastische modellen voldoende generiek zijn en daardoor in staat zijn de vloeistofstroming correct te voorspellen.

Deze dissertatie beschrijft een multilevel UQ aanpak, welke wordt toegepast op vloeistof dynamica problemen. De aandacht gaat specifiek uit naar het ontwikkelen van snelle propagatie methoden. We laten ook de brede toepasbaarheid van de stochastische modellen zien. De volgende stap is om, gebruik makende van de in dit werk ontwikkelde methodes, efficiënte UQ algoritmes te ontwikkelen om gebruik makende van data parameters te kalibreren en model onzekerheid te kwantificeren.

Contents

1

Introduction

Due to formidable progress in computers and algorithms, many problems in science and engineering can be simulated with a high level of detail using computer programs. These programs encode mathematical models describing the underlying physics of the problem. The models also contain a number of physical parameters that describe the state of the physical system. Predictions of these models may become uncertain when the true physics or parameters describing the reality are unknown or due to natural randomness in the physical system.

Sources of uncertainty in a mathematical model are typically categorized as *epistemic* uncertainty (due to lack of knowledge) or *aleatoric* uncertainty (due to inherent randomness). Within epistemic uncertainty, we encounter parametric uncertainties where the parameters that are input to the mathematical model are not completely known, for instance, the permeability field in subsurface flow modeling. Another type of epistemic uncertainty is model bias (or structural uncertainty) that arises due to inaccuracies in a mathematical model. In some cases, simplifications made just to reach stable numerical solutions may also introduce a model bias. The uncertainty in the epistemic sources can be reduced by acquiring more data regarding the model parameters or by incorporating additional physics in the mathematical model.

Some physical systems exhibit uncertainties that are irreducible i.e. cannot be reduced by acquiring more information about the process. Due to the inherent randomness, the outcome of these processes may differ even when the same experiment is repeated multiple times (for example, drawing cards from a shuffled pack). Typically, aleatoric variables are sampled from some probability distribution that is part of the model of the system.

One may also have to deal with operational uncertainties, for example, when exact boundary or initial conditions are not available. For some problems, with geometric uncertainty where the geometry of the problem is not known accurately, for example, the shape of an airfoil. The above uncertainty can again be classified as epistemic or aleatoric uncertainty. This distinction will be relevant in order to decide whether data acquisition is needed to reduce the uncertainty.

Stochastic extensions of deterministic models can be derived by modeling the inputs as random variables. For example, in subsurface flow modeling it is common to model the unknown permeability as a lognormal random field. Uncertainties due to model bias can be dealt with in a number of ways. A common practice is the addition of a discrepancy function to capture the model bias [1]. Another possibility is to perturb the quantities obtained from deterministic simulations. The amount of perturbation is typically decided based on some physical constraints. In the turbulence modeling literature, a number of stochastic turbulence models have been derived by perturbing the Reynolds stresses obtained from deterministic simulations [2–4].

Once the stochastic model is formulated, an efficient Uncertainty Quantification (UQ) technique is required to compute the statistical moments of the output quantities of interest. This choice typically depends on the uncertain dimension. Other practical factors such as the ease of implementation and the availability of an iterative solver which is robust with respect to the random input, also play a role in the selection of a suitable UQ technique. These UQ techniques are sometimes categorized as either intrusive (requires modifications in the deterministic solver) or non-intrusive (deterministic solver is used as a black-box) approaches. Notably one of the most popular intrusive techniques is the stochastic Galerkin (SG) method based on the generalized Polynomial Chaos (gPC) expansion [5–7]. The SG formulation of the stochastic PDEs typically results in a set of coupled deterministic equations. For SG methods, the existing iterative solvers cannot be directly utilized and may need drastic modifications. Furthermore, the implementation becomes highly involved when we deal with complicated PDEs involving nonlinear terms or multi-physics problems.

Another class of well-established UQ techniques includes stochastic collocation (SC) methods [7, 8] that are based on deterministic sampling approaches. The sampling nodes in the random space are computed using *cubature rules*. The main advantage of the SC method lies in its non-intrusiveness, hence, existing deterministic solvers can be utilized. The computational cost is governed by the number of nodes, that however increase exponentially with an increase in the stochastic dimensions. Another drawback of the conventional SC approach is that it may not adequately represent those regions in the stochastic space where strong nonlinearities or discontinuities may be present. Sparse grid based SC methods are nowadays standard as they allow us to reduce the number of nodes drastically without losing the order of accuracy. However, for very large uncertainty dimensions, the sparse grid SC methods can still be very expensive.

In this thesis, we will work with fluid flow problems and high-dimensional uncertainties. For such problems, Monte Carlo (MC) type methods have been proven to be effective UQ tools as they do not suffer from the curse of dimensionality, are easy to implement and have a high parallelization potential. Moreover, these MC-type methods can accurately represent the entire stochastic space given a sufficiently large number of samples. A well-known drawback of the plain MC method is the slow convergence of the sampling error, with $(\mathbb{V}[Q]/N)^{1/2}$, where $\mathbb{V}[Q]$ is the sample variance of the quantity of interest Q and N is the number of samples. This makes the plain MC method very expensive for flow problems that are already computationally intensive in their deterministic setting. We point out that quasi-Monte Carlo methods which use deterministically chosen sample points in an appropriate (usually high-dimensional) parameter space can yield a convergence rate better than $\mathcal{O}(N^{-1/2})$. However, we do not investigate QMC based estimators in this work. The focus of this thesis is to improve the efficiency of the standard MC method using a *multilevel* framework.

1.1. Multilevel solvers

A combination of multilevel approaches can, in principle, help us to eliminate the computational complexities arising due to both MC sampling as well as the spatio-temporal scales characteristics. In this thesis, we report *multilevel solvers* for fluid flow problems where complexities associated to the sampling are overcome by the multilevel Monte Carlo (MLMC) method and complexities due to spatio-temporal scales via the multigrid solver.

The idea of the MLMC method was introduced by Heinrich [9] to speed up the computation of high-dimensional integrals. This multilevel idea was further developed and improved by Giles [10] to reduce the order of complexity of Monte Carlo path simulations for stochastic differential equations (SDEs). The improved efficiency of the MLMC method comes from building the estimate for a quantity of interest Q, on a hierarchy of grids or levels, by exploiting the linearity of the expectation operator, i.e.

$$\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}[Q_\ell - Q_{\ell-1}], \qquad (1.1)$$

for all $\ell \in \{0, 1, ..., L\}$. On the coarsest grid, for $\ell = 0$, the expectation is inexpensive to compute accurately. For large values of ℓ where the numerical solution is comparatively expensive, only a few realizations are required as the variance of the correction term $\mathbb{V}[Q_\ell - Q_{\ell-1}]$ is significantly smaller compared to the variance of Q_ℓ, i.e. $\mathbb{V}[Q_\ell]$. While offering large savings over the standard MC method, MLMC retains all the important properties of MC methods like parallelization and possible combination with other variance reduction techniques (see e.g. [11–13]).

As for any non-intrusive UQ method, the computational cost of the MLMC estimator also depends on the efficiency of the black-box solver used to compute samples on different mesh resolutions. Combining the MLMC method with multigrid solvers [14–16] seems ideal as multigrid solvers can, in general, obtain scale independent convergence rates. Classical iterative methods such as Jacobi or Gauss-Seidel iterations are only efficient for eliminating high-frequency errors between the exact solution and its numerical approximation. When these high-frequency errors are projected on coarser grids they again appear oscillatory. Therefore, by using a hierarchy of grids, these basic iterative schemes can be made efficient for eliminating all high and low error components. This idea forms the basis for a multigrid solver. A multigrid iteration starts with a basic iterative relaxation on the finest grid called the *pre-smoothing*. Then, the residual obtained after smoothing is transferred to the next coarser grid using an appropriate *restriction operator*. As the same problem needs to be solved on the coarser grid, the same relaxation method may be applied to eliminate the corresponding components of the error. This process is recursively extended until the coarsest grid is reached where a direct solution of the problem is cheap. After that, the corrections obtained on the coarsest level is transferred to the next finer grid using an appropriate *prolongation operator*. Here, additional relaxations known as the *post-smoothing* may be performed to remove error components introduced from the prolongation step. This is repeated until the finest grid is reached. This completes one multigrid cycle. A number of multigrid cycles may be required to reach up to the discretization accuracy. Note that the efficiency of the multigrid solver depends on the choice of smoothers, transfer operators, coarse grid discretizations, etc.

A good initial approximation can accelerate the convergence of an iterative scheme. Obtaining a good initial approximation by means of inexpensive coarser grid computations is known as nested iteration. The combination of nested iteration and multigrid cycling results in the so-called full multigrid (FMG) algorithm. The FMG algorithm is the most efficient class of multigrid solvers since it is considered to be asymptotically optimal, i.e. the number of arithmetic operations required is proportional to the number of grid points, with only a small constant of proportionality. Hence, combining the MLMC method with the FMG solvers also open up the possibilities for designing linear-scaling UQ techniques.

Developing a multigrid method for problems with random inputs is challenging, as depending on the random inputs, the multigrid convergence rate may have a high variance or even diverge for certain samples. Therefore, another focus of this thesis is to develop multigrid solvers that are suitable for the MLMC estimator. We pay special attention to the choice of the multigrid components and provide modifications needed in the deterministic multigrid solver to deal with the stochasticity.

1.2. Stochastic flow problems

We will work with four different flow problems with applications ranging from subsurface flow to aerodynamics. They can be further categorized into linear and nonlinear PDE models; single and multi-phase flows and also based on the number of physical processes occurring simultaneously, as single-physics or multi-physics problems. Next, we discuss the relevance of the considered problems and the motivation to use multilevel methods to solve their stochastic formulations.

1.2.1   Darcy flow

Darcy's equation [17] is used to study the flow in a porous medium. This model finds its application in hydrogeology to study the groundwater flow through an aquifer. To model such flows, the permeability values should be known at all locations in the domain. This is usually not feasible and only a few sparse measurements are available. It is common practice to model the subsurface permeability as a lognormal random field [18–20]. Further, as a realistic geologic formation exhibit highly fluctuating permeabilities, its random realizations should be based on covariance functions with a small correlation length and a large variance in order to enforce a low spatial regularity in the samples of the permeability field. Such covariance models translate to a high-dimensional uncertainty as it requires a large number of Fourier modes (in the context of the Karhunen-Loève expansion) to adequately represent the random fields. Moreover, due to low spatial regularity, the PDE must be solved on a sufficiently refined mesh to reduce the discretization error up to reasonably low values. Hence, for many realistic problems, a combination of the MLMC method with multigrid solver may be the only feasible approach for UQ.

1.2.2   Transport in a Darcy-Stokes system

A multi-physics Darcy-Stokes model can be used to study the interaction between surface water and groundwater flow. This model can be employed to design and analyze filtration devices used for industrial applications, simulate blood flows, etc. The coupled Darcy-Stokes model, when combined with a transport equation, for instance, the advection-diffusion equation can be utilized to study different dynamical processes, like pollution. The combined model is highly relevant for the investigation and assessment of the accidental discharge of radioactive contaminants or chemical spills in the surface water bodies and the subsequent transport to the connected aquifers.

The transport in the coupled Darcy-Stokes flow becomes uncertain when certain physical quantities such as fluid viscosity or permeability of the porous media are not completely known. Furthermore, for some contaminant transport problems, initial and inflow conditions may also be uncertain. Uncertainty is further exacerbated if a lower-order discretization scheme is used in the case of advection-dominated flows due to excessive numerical diffusion. Typically, the computational cost for solving even the deterministic version can be very expensive. Therefore, computing stochastic moments such as the mean, variance for relevant output quantities using plain Monte Carlo or deterministic sampling methods can easily become intractable. We propose a special combination of numerical schemes for each subproblem, that, when utilized in a multilevel framework, greatly reduce the computational cost for UQ of stochastic transport.

1.2.3   Multi-phase flow in porous media

Many problems in water resource management related to irrigation systems, groundwater recharge, mining operation, etc., require a robust and accurate simulation of multi-phase (air and water) variably saturated flows. These flows are commonly simulated using the nonlinear Richards' equation [21]. This PDE model in combination with van Genuchten and Mualem parameterization are highly successful in predicting transport for a number of soil types. As the solutions are extremely sensitive to the soil parameters, the prediction capability of these models is influenced by the quality of the soil measurement data. For many problems, complete information of these quantities is not available, necessitating probabilistic modeling of these parameters. We can derive the stochastic formulation of Richards' equation by modeling these soil parameters as random fields.

The randomness in these soil parameters may result in path-wise highly nonlinear systems, therefore, Monte Carlo based methods are most appropriate for these uncertain PDEs, as they can adequately represent the complete stochastic space. Using the plain vanilla Monte Carlo method for the strongly nonlinear Richards' equation is highly expensive, as widely used solvers (based on,

for e.g., Picard or Newton iteration) require a very fine spatio-temporal mesh to converge to the correct solution. Moreover, a straightforward application of the MLMC method may not be very effective as only a few coarse grids can be utilized, due to the nonlinearity.

1.2.4   Turbulence modeling

Turbulence is a multi-scale phenomenon, which presents a challenge for simulation. Resolving all scales of turbulent flow is impractical, necessitating the modeling of turbulence effects. Reynolds-averaged Navier-Stokes (RANS) approaches split the flow into time-averaged and fluctuating parts, and the fluctuating parts are modeled. This modeling approach has historically been performed with a combination of physical insight and empirical data-fitting and has resulted in the RANS model becoming the dominant fluid modeling paradigm in engineering. However progress in RANS modeling has stalled to some extent; existing models can be inaccurate, and no error estimates exist. Uncertainty quantification has a role in continuing the improvement of RANS closure models. In particular, all models have model-form uncertainty, in which particular modeling assumptions made are known to be unreliable in some cases (often under specific, known conditions). A key example is the *Boussinesq hypothesis*, and there are many attempts to introduce uncertainty into closure models to relax this assumption. These attempts consist of introducing randomness into the modeled Reynolds Stress Tensor (RST), either by perturbing eigenvalues or tensor-invariants [2, 3, 22], or by perturbing the entire matrix [4]. Generally, these perturbations are specified globally, in order to keep the uncertain dimension moderate. This is contrary to the fact that the Boussinesq approximation is only locally an inaccurate approximation. We, therefore, need methods capable of efficient propagation of high-dimensional field uncertainty in a RANS prediction.

The MLMC approach is an attractive option for efficient propagation of uncertainty in RANS turbulence models as it can be used to design estimators for which the computational complexity is the same as solving a deterministic problem on the finest grid. Furthermore, these estimators are relatively easy to implement. There are a number of challenges that need to be overcome in order to apply MLMC to RANS problems, for instance, when the grid convergence is unreliable for the considered hierarchy of levels. Another concern is when the fine-scale features (e.g. boundary-layers) are unresolved on coarse levels and may not result in any variance reduction for the MLMC estimator.

1.3. Roadmap

The remainder of the thesis is organized as follows:

Chapter 2 contains the theory of multilevel Monte Carlo methods applied to a generic stochastic fluid flow problem. This chapter covers the core material that is utilized in the succeeding chapters. We discuss different sampling strategies along with the key factors that control the cost of the MLMC estimator. Recent advancements in the multilevel Monte Carlo method are also reviewed.

In Chapter 3, we show that a high-order discretization scheme can improve the computational complexity of the MLMC method. We demonstrate this for a Darcy flow problem, where we utilize a fourth-order finite-volume discretization to show improved MLMC convergence compared to the commonly utilized second-order schemes. The resulting fourth-order discrete system is solved using a multigrid solver based on a defect correction strategy build upon a cell-centered multigrid method. Additionally, we also discuss a practical way of combining a full multigrid solver with the multilevel Monte Carlo estimator constructed on the same mesh hierarchy. The contents of this chapter are published in [23].

In Chapter 4, we investigate an approach to quantify the statistics of multigrid solvers when utilized to solve partial differential equations with random coefficients. Specifically, our analysis

method is based on a non-standard Local Fourier Analysis (LFA) that can assess the multigrid convergence for problems with random and jumping coefficients. This LFA technique is described for a simple cell-centered multigrid method for Darcy flow problem in a random porous medium. The statistics extracted from this analysis help us to estimate apriori the time needed for solving certain uncertainty quantification problems by means of a multigrid multilevel Monte Carlo method. The contents of this chapter have been presented in [24].

Chapter 5 deals with uncertainty quantification of advection-dominated contaminant transport in a coupled Darcy-Stokes flow system using the multigrid multilevel Monte Carlo method. The uncertainty in the flow arises due to an uncertain permeability field for the porous media. We show that the contaminant transport is highly sensitive to the heterogeneity of the permeability field and also to the accuracy of the discretization scheme. To solve the steady-state Darcy-Stokes flow with a highly heterogeneous diffusion coefficient, a specific monolithic multigrid algorithm is presented. Furthermore, the unsteady advection-dominated transport equation is numerically solved using the Alternating Direction Implicit (ADI) based implicit time-stepping scheme. We demonstrate that the proposed combination of the numerical schemes is efficient in generating samples for the MLMC estimator. The findings of this work are published in [25].

In Chapter 6, we solve the stochastic nonlinear Richards' equation using a parametric continuation based multilevel Monte Carlo estimator. The key idea is to solve simpler nonlinear problems on the coarser levels and the strongly nonlinear target problem only on the finest grid level. When dealing with strongly nonlinear problems, the standard MLMC method may not be very effective, as these problems require highly refined meshes to converge, thus, only a few expensive grids can be incorporated in the MLMC hierarchy. To demonstrate the potential of our novel approach, we utilize the widely used modified Picard iteration as the solution method and show significant speed-up for a number of test problems. The contents of this chapter are presented in [26].

Chapter 7 extends the MLMC framework to stochastic turbulence modeling, for quantifying model-form uncertainties associated with the Reynolds-Averaged Navier-Stokes (RANS) simulations. We consider two high-dimensional stochastic extensions of the RANS equations to demonstrate the applicability of the MLMC method. The first extension is derived from global perturbation of the baseline eddy viscosity field using a lognormal random field. A more general second approach is considered where the Reynolds Stress Tensor is sampled as a positive definite random matrix with spatial correlation ensuring realisability. Experiments for two fundamental flows along with comparisons in terms of cost and accuracy are made with the plain Monte Carlo method. Results of this chapter are reported in [27].

Chapter 8 draws conclusions for each chapter and for the complete thesis work. We also pose some open research questions and discuss topics for further research.

2

Theory of multilevel Monte Carlo

In this chapter a detailed theoretical description of the multilevel Monte Carlo estimator is presented. We describe the method for a generic stochastic fluid flow problem and compare the standard and multilevel Monte Carlo estimators in terms of accuracy and computational cost. Different sampling strategies to determine level-dependent samples are also discussed. Also a brief summary on the recent developments in multilevel Monte Carlo methods is provided.

2.1. Introduction

Let us define a general mathematical model for stochastic flow problems

$$\mathcal{L}(\mathbf{x}, \omega, u(\mathbf{x}, \omega)) = f(\mathbf{x}, \omega), \tag{2.1}$$

where $\mathcal{L}$ denote a stochastic differential operator; $u(\mathbf{x}, \omega)$ represents solution (i.e. flow field) and $f(\mathbf{x}, \omega)$ is source. Both are functions on physical space $\mathbf{x} \in \mathcal{D} \subset \mathbb{R}^n$ and random event $\omega \in \Omega$. The complete probability space is denoted by $(\Omega, \mathbb{F}, \mathbb{P})$, where Ω is the sample space with σ-field $\mathbb{F} \subset 2^\Omega$ as a set of events and probability measure $\mathbb{P} : \mathbb{F} \to [0, 1]$. We further assume that $u(\mathbf{x}, \omega)$ belongs to the functional space $L^2(\Omega, \mathcal{D})$ corresponding to the space of square-integrable measurable functions $u : \Omega \to L^2(\mathcal{D})$ for the probability space $(\Omega, \mathbb{F}, \mathbb{P})$. These spaces are equipped with the norm

$$\left\|u(\mathbf{x}, \omega)\right\|_{L^2(\Omega, \mathcal{D})} := \mathbb{E}\left[\left\|u(\mathbf{x}, \omega)\right\|_{L^2(\mathcal{D})}^2\right]^{\frac{1}{2}} = \left(\int_\Omega \int_\mathcal{D} u(\mathbf{x}, \omega) \mathrm{d}\mathbf{x} \mathrm{d}\omega\right)^{\frac{1}{2}}. \tag{2.2}$$

The discrete version of the continuous problem (2.1) on a spatial mesh $\mathcal{D}_h$ is defined as

$$\mathcal{L}_h(\mathbf{x}_h, \omega, u_h(\mathbf{x}_h, \omega)) = f_h(\mathbf{x}_h, \omega), \tag{2.3}$$

where $\mathbf{x}_h \in \mathcal{D}_h$ with $h > 0$ as the largest mesh width in $\mathcal{D}_h$ and quantities $\mathcal{L}_h, u_h$ and f_h are the discrete approximations of $\mathcal{L}, u$ and f, respectively. We further assume that under the limit $h \to 0$, the model (2.3) converges to the exact solution, i.e. $u_h \to u$.

Our aim is to estimate the stochastic moments of some functional of the solution u_h using the multilevel Monte Carlo method. We first introduce the standard MC estimator as it forms the building block for the MLMC estimator. We will closely follow the descriptions from [10, 13, 28–32].

2.2. Standard Monte Carlo

For simplicity, we use the solution field u_h as the quantity of interest (QoI). Assuming that we are able to compute N independent identically distributed (i.i.d) random samples $\{u_h(\omega_i)\}_{i=1}^N$ on the computational grid $\mathcal{D}_h$, the unbiased MC estimator for $\mathbb{E}[u_h]$ can be derived as

$$\mathbb{E}[u_h] \approx \mathcal{E}_N^{MC}[u_h] := \frac{1}{N} \sum_{i=1}^N u_h(\omega_i), \tag{2.4}$$

where $\omega_i \in \Omega$. The above estimator is easy to implement. One needs to generate N i.i.d. samples of random inputs and solve the discrete flow problem (2.3) for each sample. These samples are averaged to obtain the MC estimate $\mathcal{E}_N^{MC}[u_h]$. Similarly, the unbiased estimator for the variance $\mathbb{V}[u_h]$ is defined as

$$\mathbb{V}[u_h] \approx \mathcal{V}_N^{MC}[u_h] := \frac{1}{N-1} \sum_{i=1}^N \left(u_h(\omega_i) - \mathcal{E}_N^{MC}[u_h] \right)^2. \tag{2.5}$$

Higher moments of u_h can be computed analogously from the ensemble.

2.2.1   Accuracy of the MC estimator

We quantify the accuracy of the MC estimator $\mathcal{E}_N^{MC}[u_h]$ in terms of the L^2- based mean square error (MSE) measure $\left\| \mathbb{E}[u] - \mathcal{E}_N^{MC}[u_h] \right\|_{L^2(\Omega,\mathcal{D})}^2$. There are two sources of error in $\mathcal{E}_N^{MC}[u_h]$ due to discretization and finite sampling. Using the triangle inequality, the two errors can be expressed separately and the MSE can be bounded as

$$\left\| \mathbb{E}[u] - \mathcal{E}_N^{MC}[u_h] \right\|_{L^2(\Omega,\mathcal{D})}^2 \leq \left\| \mathbb{E}[u] - \mathbb{E}[u_h] \right\|_{L^2(\mathcal{D})}^2 + \left\| \mathbb{E}[u_h] - \mathcal{E}_N^{MC}[u_h] \right\|_{L^2(\Omega,\mathcal{D})}^2. \tag{2.6}$$

The discretization error can simply be bounded as:

$$\left\| \mathbb{E}[u - u_h] \right\|_{L^2(\mathcal{D})} \leq c_0 h^\alpha, \tag{2.7}$$

where c_0 is a constant and $\alpha > 0$ is the rate of decay of the discretization bias. As the exact solution $\mathbb{E}[u]$ is not available, a relative error measure, for example $\left\| \mathbb{E}[u_h] - \mathbb{E}[u_{2h}] \right\|_{L^2(\mathcal{D})}$ can be used to bound the exact discretization error, as

$$\left\| \mathbb{E}[u - u_h] \right\|_{L^2(\mathcal{D})} \leq \frac{\left\| \mathbb{E}[u_h - u_{2h}] \right\|_{L^2(\mathcal{D})}}{2^\alpha - 1}. \tag{2.8}$$

The above relation can be easily derived using the reverse triangle inequality and (2.7). Here the quantity $\left\| \mathbb{E}[u_h - u_{2h}] \right\|_{L^2(\mathcal{D})}$ can be numerically approximated using the MC method with a few samples. Similarly, the rate α can also be derived numerically. Typically, the value of α depends on the regularity of the QoI in the stochastic and physical space, and also on the order of the discretization scheme used to solve the PDE.

The next task is to quantify the sampling error $\left\|\mathbb{E}[u_h] - \mathcal{E}_N^{MC}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}^2$. We can show that

$$
\begin{aligned}
\left\|\mathbb{E}[u_h] - \mathcal{E}_N^{MC}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}^2 &= \mathbb{E}\left[\left\|\mathbb{E}[u_h] - \frac{1}{N}\sum_{i=1}^{N} u_h(\omega_i)\right\|_{L^2(\mathcal{D})}^2\right] \\
&= \mathbb{E}\left[\frac{1}{N^2}\left\|\sum_{i=1}^{N}\mathbb{E}[u_h] - u_h(\omega_i)\right\|_{L^2(\mathcal{D})}^2\right] \\
&= \frac{1}{N^2}\mathbb{E}\left[\int_{\mathcal{D}}\left(\sum_{i=1}^{N}\mathbb{E}[u_h(\mathbf{x},\cdot)] - u_h(\mathbf{x},\omega_i)\right)\right. \\
&\qquad\qquad \left.\left(\sum_{j=1}^{N}\mathbb{E}[u_h(\mathbf{x},\cdot)] - u_h(\mathbf{x},\omega_j)\right)d\mathbf{x}\right].
\end{aligned}
\tag{2.9}
$$

Due to the fact that the samples are mutually independent, the terms with $i \neq j$ vanish and by linearity of the expectation, we may write

$$
\begin{aligned}
\left\|\mathbb{E}[u_h] - \mathcal{E}_N^{MC}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}^2 &= \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}\left[\int_{\mathcal{D}}\left(\mathbb{E}[u_h(\mathbf{x},\cdot)] - u_h(\mathbf{x},\omega_i)\right)^2 d\mathbf{x}\right] \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\left\|\mathbb{E}[u_h] - u_h\right\|_{L^2(\Omega,\mathcal{D})}^2 \\
&= \frac{\left\|\mathcal{V}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}}{N},
\end{aligned}
\tag{2.10}
$$

where the term $\left\|\mathcal{V}[u_h]\right\|_{L^2(\Omega,\mathcal{D})} := \left\|\mathbb{E}[u_h] - u_h\right\|_{L^2(\Omega,\mathcal{D})}^2$ denotes the sample variance and can be numerically estimated. Therefore, the MSE of the $\mathcal{E}_N^{MC}[u_h]$ can be compactly expressed as

$$
\left\|\mathbb{E}[u] - \mathcal{E}_N^{MC}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}^2 \leq (c_0 h^\alpha)^2 + \frac{\left\|\mathcal{V}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}}{N}.
\tag{2.11}
$$

For a given mesh $\mathcal{D}_h$, the MC estimator can achieve an MSE of $\mathcal{O}(h^{2\alpha})$ at best. Therefore, the number of MC samples to achieve convergence are obtained by balancing the sampling error with the discretization error, i.e.

$$
N = \mathcal{O}(h^{-2\alpha}).
\tag{2.12}
$$

More generally, to achieve an MSE of $\mathcal{O}(\varepsilon^2)$, we can choose $h = \mathcal{O}(\varepsilon^{1/\alpha})$ and $N = \mathcal{O}(\varepsilon^{-2})$, assuming $\left\|\mathcal{V}[u_h]\right\|_{L^2(\Omega,\mathcal{D})}$ is constant, independent of h. With this, if we express the cost to compute one sample of u_h by $\mathcal{W}_h = \mathcal{O}(h^{-\gamma})$, where $\gamma \geq 1$, we obtain a total cost of

$$
\mathcal{W}_{h,N}^{MC}(\varepsilon) = N\mathcal{W}_h = \mathcal{O}(\varepsilon^{-2-\gamma/\alpha}),
\tag{2.13}
$$

to achieve a MSE of $\mathcal{O}(\varepsilon^2)$. It is pointed out that the computational complexity of the standard MC method can be improved using an optimal solver, for e.g., a full multigrid (FMG) solver with $\gamma \approx d$, where d is the number of spatial dimensions. Another way to reduce the cost is by using a higher-order discretization, i.e. with an increased value for α. A larger value of α can give the same accuracy on a coarser grid and can therefore make the MC simulation significantly cheaper. In practice the plain MC method can be very expensive. Consider, for instance, a 2D problem with an optimal solver with $\gamma = 2$ and a second-order accurate discretization with $\alpha = 2$. In this case, we

get a complexity of $\mathcal{O}(\varepsilon^{-3})$, therefore to reduce the MSE by a factor of 2, we need to spend 8 times more computational effort.

Next, we show that by using a multilevel approach the complexity of the plain MC can be improved, and for some problems even an optimal multilevel estimator can be derived, in the sense that the total cost is only a small constant factor of one deterministic solve on the finest resolution.

2.3. Multilevel Monte Carlo

In the following, we describe MLMC estimators for the same quantity u_h. First, we construct a hierarchy of spatial grids $\{\mathcal{D}_\ell\}_{\ell=0}^L$ for the spatial domain $\mathcal{D}$ where the largest cell-width for level ℓ defined as

$$h_\ell = s^{-\ell} h_0, \quad \text{for} \quad h_0 > 0, \tag{2.14}$$

where h_0 is largest cell-width on the coarsest mesh $\mathcal{D}_0$ and $s > 1$ represents a grid refinement factor. Now, using the linearity of the expectation operator, we can build the estimate on this hierarchy of meshes as

$$\mathbb{E}[u_{h_L}] = \mathbb{E}[u_{h_0}] + \sum_{\ell=1}^L \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}]. \tag{2.15}$$

On the coarsest grid for $\ell = 0$, expectations are inexpensive to compute accurately and for large values of ℓ, where numerical solutions are comparatively expensive, only a few realizations are required as the variance of the correction term $\mathbb{V}[u_{h_\ell} - u_{h_{\ell-1}}]$ is significantly smaller compared to the variance of u_{h_ℓ}, i.e. $\mathbb{V}[u_{h_\ell}]$. Our description of the MLMC estimator will be based on a geometric hierarchy of grids, however this is not a necessary criterion. A valid MLMC estimator can be constructed using a hierarchy that consists of levels with increasing cost and accuracy.

Each of the expectations in (2.15) can be approximated using the standard MC estimator defined in (2.4) as

$$
\begin{aligned}
\mathbb{E}[u_{h_L}] \approx \mathcal{E}_L^{ML}[u_{h_L}] &:= \sum_{\ell=0}^L \mathcal{E}_{N_\ell}^{MC}[u_{h_\ell} - u_{h_{\ell-1}}] \\
&= \sum_{\ell=0}^L \left(\frac{1}{N_\ell} \sum_{i=1}^{N_\ell} u_{h_\ell}(\omega_i) - u_{h_{\ell-1}}(\omega_i) \right),
\end{aligned}
\tag{2.16}
$$

with $N_\ell \in \mathbb{N}$ denoting the level-dependent number of MC samples. For notational convenience, we set $u_{h_{-1}} = 0$. The MC estimators at different levels are mutually independent. The samples $u_{h_\ell}(\omega_i) - u_{h_{\ell-1}}(\omega_i)$ at any level ℓ on two discretization levels h_ℓ and $h_{\ell-1}$ are generated using the same random input ω_i. This is done in order to keep the sampling variance $\mathbb{V}[u_{h_\ell} - u_{h_{\ell-1}}]$ small, thus, only a few samples are required to compute $\mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}]$ accurately.

2.3.1 Accuracy of the MLMC estimator

Similar to the MC estimator, the MSE in the MLMC estimator $\mathcal{E}_L^{ML}[u_{h_L}]$ can be represented as

$$\left\| \mathbb{E}[u] - \mathcal{E}_L^{ML}[u_{h_L}] \right\|_{L^2(\Omega,\mathcal{D})}^2 \leq \left\| \mathbb{E}[u] - \mathbb{E}[u_{h_L}] \right\|_{L^2(\mathcal{D})}^2 + \left\| \mathbb{E}[u_{h_L}] - \mathcal{E}_L^{ML}[u_{h_L}] \right\|_{L^2(\Omega,\mathcal{D})}^2. \tag{2.17}$$

The first term corresponds to the discretization error and is bounded in a similar manner as in (2.8), that is,

$$\left\| \mathbb{E}[u - u_{h_L}] \right\|_{L^2(\mathcal{D})} \leq c_0 h_L^\alpha. \tag{2.18}$$

The second term in (2.17) corresponds to the sampling error of the MLMC estimator. Using similar arguments as in (2.9)-(2.10), we can show that

$$
\begin{aligned}
\left\| \mathbb{E}[u_{h_L}] - \mathcal{E}_L^{ML}[u_{h_L}] \right\|_{L^2(\Omega,\mathcal{D})}^2 &= \mathbb{E}\left[\left\| \sum_{\ell=0}^{L} \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}] - \mathcal{E}_{N_\ell}^{MC}[u_{h_\ell} - u_{h_{\ell-1}}] \right\|_{L^2(\mathcal{D})}^2 \right] \\
&= \mathbb{E}\sum_{\ell=0}^{L}\left[\left\| \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}] - \mathcal{E}_{N_\ell}^{MC}[u_{h_\ell} - u_{h_{\ell-1}}] \right\|_{L^2(\mathcal{D})}^2 \right] \\
&= \sum_{\ell=0}^{L} \mathbb{E}\left[\left\| \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}] - \mathcal{E}_{N_\ell}^{MC}[u_{h_\ell} - u_{h_{\ell-1}}] \right\|_{L^2(\mathcal{D})}^2 \right] \\
&= \sum_{\ell=0}^{L}\left[\frac{1}{N_\ell^2}\sum_{i=1}^{N_\ell} \mathbb{E}\left[\left\| \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}] - [u_{h_\ell}(\omega_i) - u_{h_{\ell-1}}(\omega_i)] \right\|_{L^2(\mathcal{D})}^2 \right] \right] \\
&= \sum_{\ell=0}^{L} \frac{\|\mathcal{V}_\ell\|_{L^2(\Omega,\mathcal{D})}}{N_\ell},
\end{aligned}
\tag{2.19}
$$

where $\|\mathcal{V}_\ell\|_{L^2(\Omega,\mathcal{D})} := \left\| \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}] - [u_{h_\ell} - u_{h_{\ell-1}}] \right\|_{L^2(\Omega,\mathcal{D})}^2$ corresponds to the level-dependent variance. Thus, the sampling error of the MLMC estimator is just the sum of sampling errors from individual MC estimators in (2.16). We further assume that the level-dependent variance decays with a rate β as

$$
\|\mathcal{V}_\ell\|_{L^2(\Omega,\mathcal{D})} = \mathcal{O}(h_\ell^\beta) \quad \ell = 1, 2, .., L, \quad \beta > 0.
\tag{2.20}
$$

Now, the MSE of the MLMC estimator can be expressed as

$$
\left\| \mathbb{E}[u] - \mathcal{E}_L^{ML}[u_{h_L}] \right\|_{L^2(\Omega,\mathcal{D})}^2 \leq \left(c_0 h_L^\alpha\right)^2 + \sum_{\ell=0}^{L} \frac{\|\mathcal{V}_\ell\|_{L^2(\Omega,\mathcal{D})}}{N_\ell}.
\tag{2.21}
$$

Again, to obtain a tolerance of $\mathcal{O}(\varepsilon^2)$ both errors in the above bound should be reduced to $\mathcal{O}(\varepsilon^2)$. The finest level grid size is obtained in the same way as the standard MC, i.e. $h_L = \mathcal{O}(\varepsilon^{1/\alpha})$. Practical implementation of the MLMC estimator however does not require a value of L in advance (see Algorithm 1). Usually, we start with a small L and increase the number of levels until the following criterion is met,

$$
\left\| \mathbb{E}[u_{h_L} - u_{h_{L-1}}] \right\|_{L^2(\mathcal{D})} < (s^\alpha - 1)\varepsilon,
\tag{2.22}
$$

which is obtained by (2.8) and taking $\left\| \mathbb{E}[u - u_{h_L}] \right\|_{L^2(\mathcal{D})} = \varepsilon$. To obtain level-dependent samples N_ℓ, a number of strategies exist in the MLMC literature. We discuss two commonly used sampling approaches that will be utilized in this thesis.

2.3.2 Sampling based on optimization

As proposed in [10, 28], the level-dependent samples N_ℓ can be obtained by solving an optimization problem that minimizes the total cost needed for reducing the sampling error to ε^2 i.e.

$$
\min\left(\sum_{\ell=0}^{L} N_\ell \mathcal{W}_\ell \right) \quad \text{s.t.} \quad \sum_{\ell=0}^{L} \frac{\|\mathcal{V}_\ell\|_{L^2(\Omega,\mathcal{D})}}{N_\ell} = \varepsilon^2,
\tag{2.23}
$$

where $\mathcal{W}_\ell = \mathcal{O}(h_\ell^{-\gamma})$ is the cost of one sample on level ℓ. By treating N_ℓ as a continuous variable and using a Lagrange multiplier ζ^2, we can formulate the constraint equation as:

$$\frac{\partial}{\partial N_\ell} \left(\sum_{k=0}^{L} N_k \mathcal{W}_k + \zeta^2 N_k^{-1} \left\| \mathcal{V}_k \right\|_{L^2(\Omega, \mathcal{D})} \right) = 0, \tag{2.24}$$

which yields

$$N_\ell = \zeta \sqrt{\frac{\left\| \mathcal{V}_\ell \right\|_{L^2(\Omega, \mathcal{D})}}{\mathcal{W}_\ell}}, \tag{2.25}$$

where $\zeta = \varepsilon^{-2} \sum_{k=0}^{L} \sqrt{\left\| \mathcal{V}_k \right\|_{L^2(\Omega, \mathcal{D})} \mathcal{W}_k}$. It is straightforward to see that for $\ell > 0$ the number of samples decays as $N_{\ell+1} = N_\ell s^{-(\beta+\gamma)/2}$ where s is the grid scaling factor defined earlier. The total cost to obtain a MSE of $\mathcal{O}(\varepsilon^2)$ is given by

$$\mathcal{W}_L^{ML}(\varepsilon) = \sum_{\ell=0}^{L} N_\ell \mathcal{W}_\ell = \varepsilon^{-2} \left(\sum_{\ell=0}^{L} \sqrt{\left\| \mathcal{V}_\ell \right\|_{L^2(\Omega, \mathcal{D})} \mathcal{W}_\ell} \right)^2. \tag{2.26}$$

Note that the cost from level ℓ is determined by the product $\left\| \mathcal{V}_\ell \right\|_{L^2(\Omega, \mathcal{D})} \mathcal{W}_\ell = \mathcal{O}(h_\ell^{(\beta-\gamma)})$. Thus, we have the following three scenarios: i) For $\beta > \gamma$, the dominant cost comes from the coarsest level ii) for $\beta = \gamma$ all levels equally contribute to the total cost and iii) for $\beta < \gamma$, the dominant cost comes from the finest level. For completeness, we now state the generalized MLMC complexity (for proof see [28]).

$$\mathcal{W}_L^{ML}(\varepsilon) = \begin{cases} \mathcal{O}\left(\varepsilon^{-2}\right), & \text{if} \quad \beta > \gamma, \\ \mathcal{O}\left(\varepsilon^{-2}(\log \varepsilon)^2\right), & \text{if} \quad \beta = \gamma, \\ \mathcal{O}\left(\varepsilon^{-2-(\gamma-\beta)/\alpha}\right), & \text{if} \quad \beta < \gamma. \end{cases} \tag{2.27}$$

Further, if $\beta < \gamma$ and $\beta = 2\alpha$, we obtain the cost $\mathcal{W}_L^{ML}(\varepsilon) = \mathcal{O}(\varepsilon^{-\gamma/\alpha}) = \mathcal{O}(h_L^{-\gamma})$, which is of the same order as the cost of one fine grid solve. For this case, the MLMC estimator is optimal as the cost has the same computational complexity as the deterministic version of the problem.

The parameters α and β depend on the regularity of the solution as well as the numerical method used to approximate the PDE and can be increased by using high-order discretizations for sufficiently smooth solutions. Typically, these parameters do not change with the spatial dimension. On the other hand, the rate γ increases with dimension and when using an optimal solver it is approximately equal to the spatial dimension d. This implies that for some quantities of interest the asymptotic cost might change when dealing with higher-dimensional problems. Therefore, a high-order MLMC scheme may help in countering this effect. Even for the case where we do not have the scope of asymptotic improvement, we can reduce the number of MLMC levels by increasing α and the number of samples by improving the parameter β.

The MLMC estimator $\mathcal{E}_L^{ML}[u_{h_L}]$ (2.16) can be computed using the following heuristic algorithm proposed by Giles [10]:

Algorithm 1 MLMC algorithm

1: Fix the tolerance ε, number of levels $L = 2$ and initial number of samples N_ℓ for $\ell = 0, 1, 2$.
2: Compute quantities $\mathcal{E}_{N_\ell}^{MC}[u_{h_\ell} - u_{h_{\ell-1}}], \left\| \mathcal{V}_\ell \right\|_{L^2(\Omega, \mathcal{D})}, \alpha, \beta$ for all levels.
3: Using (2.25) update the number of samples N_ℓ for all levels.
4: Evaluate the extra samples and update $\mathcal{E}_{N_\ell}^{MC}[u_{h_\ell} - u_{h_{\ell-1}}], \left\| \mathcal{V}_\ell \right\|_{L^2(\Omega, \mathcal{D})}, \alpha, \beta$ for all levels.
5: Test for convergence using the criterion (2.22), if converged quit and assemble $\mathcal{E}_L^{ML}[u_{h_L}]$.
6: If not converged, set $L = L + 1$ and $\left\| \mathcal{V}_L \right\|_{L^2(\Omega, \mathcal{D})} = \left\| \mathcal{V}_{L-1} \right\|_{L^2(\Omega, \mathcal{D})} s^{-\beta}$, and repeat steps 3-6.

The above algorithm is sensitive to the choice of initial number of samples. Using a small initial number of samples may lead to an abrupt termination without reaching the specified tolerance. On the other hand, using a large number of initial samples may lead to over sampling. A more detailed analysis of this algorithm is provided in [29].

2.3.3   Fixed sampling approach

In many applications, only a certain fixed number is feasible as the choice of the finest level sample N_L. In such cases, the number of samples on coarser levels can simply be derived by equilibrating sampling errors on coarser levels with the finest level, i.e.

$$\frac{||\mathcal{V}_\ell||_{L^2(\Omega,\mathcal{D})}}{N_\ell} = \frac{||\mathcal{V}_L||_{L^2(\Omega,\mathcal{D})}}{N_L} \quad \text{for} \quad \ell = 0, 1, ..., L-1. \tag{2.28}$$

Using (2.20), we obtain the sampling sequence as

$$N_\ell = \lceil N_L s^{\beta(L-\ell)} \rceil. \tag{2.29}$$

Moreover, to reduce the sampling error to the discretization error, one can choose N_L such that $||\mathcal{V}_L||_{L^2(\Omega,\mathcal{D})} / N_L = \mathcal{O}(h_L^{2\alpha})$. The authors in [30–32] have shown that this sampling strategy has a similar asymptotic cost as the sampling strategy based on the optimization described in Section 2.3.2.

This sampling approach greatly simplifies the MLMC implementation as the number of samples on all levels is known apriori. Also, the computational work can be easily distributed in a parallel computing environment. On the other hand, parallelization of the heuristic Algorithm 1 is non-trivial as every refinement step may result in a different number of extra samples on all levels.

2.4. Recent developments

Since the MLMC field is undergoing rapid development, by no means does this section serve as a comprehensive review. Only a few notable works are mentioned. A number of extensions and generalizations of the multilevel Monte Carlo method has been proposed recently. An important extension is the Multilevel Quasi-Monte Carlo (MLQMC) estimator proposed in [13, 33] for lognormal diffusion problems. For sufficiently smooth diffusion coefficients, the MLQMC estimator can result in an asymptotic cost which is better than $\mathcal{O}(\varepsilon^{-2})$. Another extension known as the Multi-index Monte Carlo (MIMC) was proposed in [34]. The approach generalizes the notion of levels to multiple dimensions. Typically, the standard MLMC estimator is based just on the hierarchy of spatial grids. Suppose for a problem, in addition to spatial refinement, the accuracy also improves with temporal refinement or with the increase in the number of basis functions used in the stochastic approximation. For such problems, a multi-dimensional hierarchy of indices can be constructed for a more effective variance reduction. Another interesting extension is represented by the unbiased variants of the multilevel estimator [35, 36]. These methods are based on *randomization of levels*, i.e. defining a probability distribution for sampling levels. These estimators are difficult to implement and the practical benefits are problem dependent.

2.5. Conclusions

In this chapter, we presented the standard MLMC method for a general stochastic fluid flow problem. In the upcoming chapters, we apply MLMC methodologies to an array of fluid flow problems. We will present different numerical strategies to improve the MLMC parameters, for instance, high-order schemes to increase α and β, and multigrid solvers tailored specifically to the problem at hand, thus improving the parameter γ. Another principle issue in designing the MLMC estimator is

the correct treatment of random inputs on two levels such that the telescopic identity is not violated. These issues will also be studied in detail in the upcoming chapters.

3

Multigrid MLMC and improvements from high-order schemes

This chapter explores high-order schemes for improving the convergence of a multilevel Monte Carlo method for elliptic partial differential equations with lognormal random coefficients, in combination with the multigrid solution method. In detail, a fourth-order accurate finite-volume discretization is discussed. With the help of the Matérn family of covariance functions, coefficient fields with different degrees of smoothness are simulated. The idea behind using a fourth-order scheme is to capture the additional regularity in the solution introduced due to higher smoothness of the random field. Second-order schemes commonly utilized for these types of problems are not able to fully exploit this additional regularity. This chapter also introduces a practical way of combining a full multigrid solver with the multilevel Monte Carlo estimator constructed on the same mesh hierarchy. Through this integration, one full multigrid solve at any level provides a valid sample for all the preceding Monte Carlo levels. A number of numerical experiments are presented confirming excellent performance of the fourth-order multilevel estimator compared to the second-order variant.

3.1. Introduction

We focus on the development and analysis of a multilevel Monte Carlo method based on a high-order finite-volume (FV) approximation of a class of elliptic PDEs with random coefficients, stated as:

$$-\nabla \cdot k\nabla p = f, \tag{3.1}$$

where the unknown p is a scalar-valued function, k is a *random field* and f is a source function, defined in some domain $\mathcal{D} \in \mathbb{R}^n$. The above PDE finds its application, for instance, in hydrogeology for modeling subsurface flows with k representing the permeability. We are particularly interested in computing expected values of different functionals of the solution p, denoted by $Q := Q(p)$. Due to the high-dimensional nature of the random field k, it becomes very challenging to obtain reliable estimates of these hydrogeological quantities. The choice of appropriate uncertainty quantification tools boils down to computational efficiency. In the present context, Monte Carlo (MC) type methods are sometimes favored [37] due to their dimension independent convergence properties and simplicity of implementation. For standard MC methods, the root-mean-square error (RMSE)

---

converges as $(\mathbb{V}[Q]/N)^{1/2}$, where N is the number of samples and $\mathbb{V}[Q]$ is the sample variance. This slow convergence with respect to N is, however, the main drawback of the method. To remedy this, various sampling and variance reduction techniques have been applied, for example, the authors in [38] have applied a quasi-Monte Carlo method to improve the convergence rate for this problem. More recently, MLMC methods have also been formulated for this problem (see e.g. [39–41]), which lead to a dramatic cost reduction compared to the classical MC approach.

There are two objectives of this chapter. First is to show that a high-order discretization scheme can be used to reduce the computational costs of the MLMC estimator for problems that exhibit high spatial smoothness. In previous work, Giles [42] has shown an improvement from $\mathcal{O}(\varepsilon^{-2}(\log \varepsilon)^2)$ complexity to $\mathcal{O}(\varepsilon^{-2})$ to achieve a RMSE of $\mathcal{O}(\varepsilon)$ using a Milstein discretization compared to an Euler path discretization for SDEs and certain financial payoffs with Lipschitz bound. This gain was achieved due to an improvement in the strong convergence order of the schemes that is central to the efficiency of the multilevel method. In this work, we choose to use a fourth-order accurate finite-volume (FV) method for solving (3.1). The key ingredient for achieving this high-order accuracy is using a fourth-order accurate quadrature rule to approximate the boundary fluxes at each control volume. So far, the MLMC literature on Darcy flow problems relies on a second-order accurate finite-element (FE) or FV discretization scheme. We show that for certain linear functionals of p the overall asymptotic cost of the MLMC method can be improved under some smoothness assumptions. Even for the case when no asymptotic gain is possible, we show a reduction in terms of the number of MLMC levels and samples.

The second objective is to define a structure in which the multilevel estimator is integrated into a multilevel solver. To demonstrate this, we use a full multigrid (FMG) solver [14] based on the same grid hierarchy as the MLMC estimator for the numerical PDE solution. We will describe in detail the modifications required to transform this solver for the MLMC method so that the telescopic identity to approximate the expectation is not violated. This method is particularly effective for quantities with low spatial regularity. Further, we also provide an efficient and scalable multigrid solver for the fourth-order linear system in 2D obtained by combining a cell-centered multigrid for the second-order discretization with the defect correction strategy.

The outline of this chapter is as follows. We begin by defining the stochastic Darcy flow problem in Section 3.2. In Section 3.3, the second- and fourth-order discretization schemes in two dimensions are described in detail. Section 3.4 is devoted to the technical details required to construct the FMG solver for both discretization schemes. In Section 3.5 we explain the procedure for coupling a FMG solver with the multilevel estimator. Sections 3.6 and 3.7 describe methods for sampling and upscaling of Gaussian random fields, respectively. In Section 3.8 we provide some numerical results for different quantities of interest and compare the cost of different MC estimators.

3.2. Stochastic Darcy flow

We consider a steady-state single phase flow in a 2D porous media

$$-\nabla \cdot (k(\mathbf{x}, \omega)\nabla p(\mathbf{x}, \omega)) = f(\mathbf{x}), \quad \text{for} \quad \mathbf{x} \in \mathcal{D} \subset \mathbb{R}^2. \tag{3.2}$$

We denote by $\omega \in \Omega$, a random event in the complete probability space $(\Omega, \mathbb{F}, \mathbb{P})$ defined earlier. Quantity p represents the fluid pressure, k is the random permeability field and f is a known source term. The PDE (3.2) is the result of the coupling between the Darcy flux $\mathbf{q} := -k\nabla p$ and the incompressiblity condition $\nabla \cdot \mathbf{q} = f$. We consider deterministic mixed Dirichlet-Neumann boundary conditions,

$$p(\mathbf{x}, \cdot) = g_D(\mathbf{x}) \quad \text{for} \quad \mathbf{x} \in \partial \mathcal{D}_D, \quad \text{and} \quad k(\mathbf{x}, \omega)\frac{\partial p}{\partial \mathbf{n}}(\mathbf{x}, \cdot) = g_N(\mathbf{x}) \quad \text{for} \quad \mathbf{x} \in \partial \mathcal{D}_N, \tag{3.3}$$

where $\partial \mathcal{D}_D$ and $\partial \mathcal{D}_N$ represent the boundaries for Dirichlet and Neumann boundary conditions, respectively, and $\mathbf{n}$ denotes the outward normal to $\partial \mathcal{D}_N$. In the context of fluid flow problems,

Dirichlet boundary conditions represent the pressure values at inflow and outflow heads, whereas the Neumann boundary conditions define the pressure gradient at the boundaries perpendicular to the main flow direction.

In stochastic subsurface flow modeling, it is well recognized that a lognormal random field may accurately represent the permeability of a naturally occurring heterogeneous porous medium [18–20]. We model the permeability field by means of a lognormal field, i.e. $\log k(\mathbf{x}, \omega) := Z(\mathbf{x}, \omega)$ is a zero mean Gaussian random field for $\mathbf{x} \in \overline{\mathcal{D}} := \mathcal{D} \bigcup \partial \mathcal{D}_D \bigcup \partial \mathcal{D}_N$ and $\omega \in \Omega$. Therefore,

$$\begin{cases} \mathbb{E}[Z(\mathbf{x}, \cdot)] = 0, \\ \mathbf{cov}(Z(\mathbf{x_1}, \cdot), Z(\mathbf{x_2}, \cdot)) = \mathbb{E}[Z(\mathbf{x_1}, \cdot)Z(\mathbf{x_2}, \cdot)], \qquad \mathbf{x_1}, \mathbf{x_2} \in \overline{\mathcal{D}}. \end{cases} \tag{3.4}$$

The lognormal property ensures a positive permeability throughout the domain. For further simplification, we consider an *isotropic* and *stationary* Gaussian process, which can be obtained from a *homogeneous* covariance function $C_\Phi : \mathbb{R}^2 \to \mathbb{R}_+$ such that

$$\mathbf{cov}(Z(\mathbf{x_1}, \cdot), Z(\mathbf{x_2}, \cdot)) = C_\Phi(r), \quad \text{with} \quad r = ||\mathbf{x_1} - \mathbf{x_2}||_2. \tag{3.5}$$

For the problem to be well-posed, we assume

$$k_{\min} := \min k(\mathbf{x}, \omega) > 0 \quad \text{and} \quad k_{\max} := \max k(\mathbf{x}, \omega) < \infty. \tag{3.6}$$

To generate samples of the Gaussian random field Z, the covariance matrix is derived from the family of *Matérn functions* [43, 44] characterized by the parameter set $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$, and has the standard form

$$C_\Phi(r) = \sigma_c^2 \frac{2^{1-\nu_c}}{\Gamma(\nu_c)} \left(2\sqrt{\nu_c} \frac{r}{\lambda_c} \right)^{\nu_c} K_{\nu_c} \left(2\sqrt{\nu_c} \frac{r}{\lambda_c} \right). \tag{3.7}$$

Here, Γ is the gamma function and K_{ν_c} is the modified Bessel function of the second kind. The parameter $\nu_c \geq 1/2$ defines the smoothness, $\sigma_c^2 > 0$ is the variance and $\lambda_c > 0$ is the correlation length of the Gaussian process. The Matérn model has great flexibility in modeling spatial processes because of parameter ν_c, which governs the differentiability of the random field. For $\nu_c = 1/2$, the Matérn function corresponds to an exponential model and for $\nu_c \to \infty$ to a Gaussian model. Furthermore, the other two parameters λ_c and σ_c^2 dictate the number of peaks and the amplitude of the random field, respectively. The realizations of Gaussian random fields are Hölder continuous i.e. $Z(\mathbf{x}, \cdot) \in C^\eta(\overline{\mathcal{D}})$ almost surely with the exponent $0 < \eta < \nu_c$ [45]. For $\nu_c > 1$, the realizations are continuously differentiable. In the recent years, (3.2) has been extensively studied and the theoretical error estimates of the solution are refined; taking into account the regularity of the lognormal coefficient field, see for e.g. [11, 39, 46–48].

3.3. Finite-volume discretization

When modeling the subsurface flow process, FV type methods are usually preferred due to the local conservation property [49]. In this section, we derive an FV approximation of the problem (3.2) using second- and fourth-order discretization schemes on a 2D cell-centered grid. Our method will follow ideas from [50–53].

We begin by defining a *cell-centered grid* $\mathcal{D}_h$ on a unit square domain as a family of nodes given by

$$\mathcal{D}_h = \left\{ (x_i, y_j) : x_i = \left(i - \frac{1}{2} \right) h, y_j = \left(j - \frac{1}{2} \right) h; i, j = 1, 2, ..., m; m = \frac{1}{h} \right\}, \tag{3.8}$$

where h is the mesh width. A control volume with center (x_i, y_j) is denoted by $\mathcal{D}_h^{(i,j)}$ and its boundary as $\partial \mathcal{D}_h^{(i,j)} = \bigcup_{\kappa=1}^4 \partial \mathcal{D}_h^{(i,j,\kappa)}$ with κ denoting the four cell faces. In each of the finite volumes $\mathcal{D}_h^{(i,j)}$,

the integral formulation of (3.2) takes the form

$$\int_{\mathcal{D}_h^{(i,j)}} -\nabla \cdot (k\nabla p)d\mathbf{x} = \int_{\mathcal{D}_h^{(i,j)}} f d\mathbf{x}. \tag{3.9}$$

Using the Gauss divergence theorem, the left-hand side of the above integral is reformulated as a boundary integral for boundary $\partial\mathcal{D}_h^{(i,j,\kappa)}$ and $\mathbf{n}_\kappa$ is the unit normal vector to that face

$$\int_{\mathcal{D}_h^{(i,j)}} -\nabla \cdot (k\nabla p)d\mathbf{x} = -\sum_{\kappa=1}^{4} \int_{\partial\mathcal{D}_h^{(i,j,\kappa)}} k\nabla p \cdot \mathbf{n}_\kappa dS_\kappa. \tag{3.10}$$

For cell ABCD in Figure 3.1, the boundary integrals are given by

$$-\sum_{\kappa=1}^{4} \int_{\partial\mathcal{D}_h^{(i,j,\kappa)}} k\nabla p \cdot \mathbf{n}_\kappa dS_\kappa = \int_A^B k\frac{\partial p}{\partial y}dx - \int_B^C k\frac{\partial p}{\partial x}dy - \int_C^D k\frac{\partial p}{\partial y}dx + \int_D^A k\frac{\partial p}{\partial x}dy. \tag{3.11}$$

We define the normal flux as $F^\kappa = k\nabla p \cdot \mathbf{n}_\kappa$ and the so-called face-averaged normal flux as $\langle F^\kappa \rangle = \frac{1}{h}\int_{\partial\mathcal{D}_h^{(i,j,\kappa)}} F^\kappa dS_\kappa$. Now, we can rewrite the integral form of the PDE in terms of face-averaged fluxes at face centers of cell ABCD,

$$h\left[\langle F^{AB} \rangle - \langle F^{BC} \rangle - \langle F^{CD} \rangle + \langle F^{DA} \rangle\right] = \int_{\mathcal{D}_h^{(i,j)}} f d\mathbf{x}. \tag{3.12}$$

We approximate the above *flux-balance* equation to second- and fourth-order accuracy in the Sections 3.3.1 and 3.3.2, respectively.
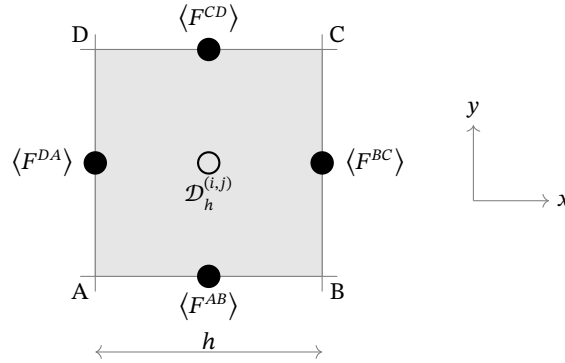
Figure 3.1: A volume $\mathcal{D}_h^{(i,j)}$ and face-averaged normal fluxes at face centers.

3.3.1   Second-order discretization

To approximate (3.12) to second-order accuracy, we use a central difference scheme to compute the gradient $\nabla p \cdot \mathbf{n}_\kappa$. For the face "BC", for example, we find

$$\langle F^{BC} \rangle^{(2)} = \frac{1}{h}\int_B^C k\frac{\partial p}{\partial x}dy = \frac{1}{h}\left[k_{i+\frac{1}{2},j}\left(\frac{p_{i+1,j} - p_{i,j}}{h}\right)h\right] + \mathcal{O}(h). \tag{3.13}$$

Fluxes at other faces are approximated similarly. To approximate the above fluxes, we need the value of coefficient along the interface. As shown in [54], a harmonic average of $k_{i,j}$ and $k_{i+1,j}$ is more general and works well for both smooth and discontinuous fields. As we only work with smooth

coefficient fields in this chapter, we approximate $k_{i+1/2,j}$ by the arithmetic mean of $k_{i,j}$ and $k_{i+1,j}$. The right-hand integral in (3.12) is approximated using the midpoint rule $\int_{\mathcal{D}_h^{(i,j)}} f \, d\mathbf{x} = h^2 f_{i,j} + \mathcal{O}(h^2)$.

Finally, we get the discrete equation for volume $\mathcal{D}_h^{(i,j)}$ as

$$-k_{i-\frac{1}{2},j} p_{i-1,j} - k_{i+\frac{1}{2},j} p_{i+1,j} + \overline{k}_{i,j} p_{i,j} - k_{i,j-\frac{1}{2}} p_{i,j-1} - k_{i,j+\frac{1}{2}} p_{i,j+1} = h^2 f_{i,j}, \tag{3.14}$$

where $\overline{k}_{i,j} = k_{i-1/2,j} + k_{i+1/2,j} + k_{i,j-1/2} + k_{i,j+1/2}$. This is a standard 5-point stencil. For Dirichlet boundary conditions, i.e. $p = g_D$, a one-sided difference instead of central difference in (3.13) is employed. A Neumann boundary condition is applied by directly using g_N in place of the finite-difference approximation.

3.3.2 Fourth-order discretization

To explain the fourth-order discretization, we shall weaken the regularity assumptions for the coefficient field and the source term. As our approximations of fluxes and the right-hand side in (3.12) are based on Taylor's expansion, we assume f and k are at least C^4, which guarantees $p \in C^5$. Since our problem involves a Hölder continuous permeability field, the convergence of the FV error for different smoothness values ν_c will be studied numerically.

We describe a fourth-order scheme for a regular 2D spatial grid. The following approach can be extended to other more complex grid systems such as mapped coordinates and locally-refined grids (see [50, 51]). To compute the face-averaged normal fluxes defined in (3.12), we use a fourth-order accurate quadrature rule. We will now explain in detail the computation of the face-averaged normal flux for the face "BC":

$$\left\langle F^{BC} \right\rangle^{(4)} = F^{BC}_{i+\frac{1}{2},j} + \frac{h^2}{24} \left. \frac{\partial^2 F^{BC}}{\partial y^2} \right|_{i+\frac{1}{2},j} + \mathcal{O}(h^4), \tag{3.15}$$

where $\left. \frac{\partial^2 F^{BC}}{\partial y^2} \right|_{i+1/2,j}$ is the transverse Laplacian of the flux at the center of face. The above relation can be derived using Taylor's expansion of the flux integrals in (3.11) [53]. We can reduce the above expression to

$$\left\langle F^{BC} \right\rangle^{(4)}_{i+\frac{1}{2},j} = \left\langle k \right\rangle_{i+\frac{1}{2},j} \left\langle \frac{\partial p}{\partial x} \right\rangle_{i+\frac{1}{2},j} + \frac{h^2}{12} \frac{\partial k}{\partial y} \left. \frac{\partial^2 p}{\partial y \partial x} \right|_{i+\frac{1}{2},j} + \mathcal{O}(h^4). \tag{3.16}$$

The above form has a smaller stencil size as (3.15). The derivation is provided in Appendix 3.A. Each term in the above expression is computed with fourth-order accuracy.

Computation of $\left\langle \frac{\partial p}{\partial x} \right\rangle_{i+\frac{1}{2},j}$

Using the same relation as (3.15), one can write

$$\left\langle \frac{\partial p}{\partial x} \right\rangle_{i+\frac{1}{2},j} = \left[\frac{\partial p}{\partial x} + \frac{h^2}{24} \frac{\partial^3 p}{\partial y^2 \partial x} \right]_{i+\frac{1}{2},j} + \mathcal{O}(h^4). \tag{3.17}$$

Next, using Taylor's expansion, we define an auxiliary variable

$$\overline{p}_{i+\frac{1}{2},j} = \frac{1}{24} \left[27(p_{i+1,j} - p_{i,j}) - (p_{i+2,j} - p_{i-1,j}) \right], \tag{3.18}$$

see Figure 3.2 (*left*) for the stencil. We can now use $\overline{\rho}_{i+\frac{1}{2},j}$ for computing the elements in (3.17) as

$$\left.\frac{\partial p}{\partial x}\right|_{i+\frac{1}{2},j} = \frac{\overline{\rho}_{i+\frac{1}{2},j}}{h} + \mathcal{O}(h^4), \tag{3.19}$$

$$\left.\frac{\partial^3 p}{\partial y^2 \partial x}\right|_{i+\frac{1}{2},j} = \frac{1}{h^3}\left[\overline{\rho}_{i+\frac{1}{2},j-1} - 2\overline{\rho}_{i+\frac{1}{2},j} + \overline{\rho}_{i+\frac{1}{2},j+1}\right] + \mathcal{O}(h^2). \tag{3.20}$$

Computation of $\left.\frac{\partial^2 p}{\partial y \partial x}\right|_{i+\frac{1}{2},j}$

For this term, we use

$$\left.\frac{\partial^2 p}{\partial y \partial x}\right|_{i+\frac{1}{2},j} = \frac{1}{2h^2}\left[\overline{\rho}_{i+\frac{1}{2},j+1} - \overline{\rho}_{i+\frac{1}{2},j-1}\right] + \mathcal{O}(h^2). \tag{3.21}$$

Derivations for the rest of the terms are provided in Appendix 3.A. This completes the fourth-order accurate face-averaged flux computation for face "BC". Averaged fluxes at other faces are computed analogously. Finally, we obtain a 21-point stencil in a 5×5 block, which is centered at the cell on which the flux divergence is computed. In case of 1D problems, we do not have any transverse derivatives and we obtain a 5-point stencil. Regarding the right-hand side of (3.12), we use the fourth-order quadrature rule on the cell $\mathcal{D}_h^{(i,j)}$, i.e.

$$\int_{\mathcal{D}_h^{(i,j)}} f d\mathbf{x} = h^2 \left[f + \frac{h^2}{24}\Delta f\right]_{i,j} + \mathcal{O}(h^4) \tag{3.22}$$

where Δ is the Laplacian operator computed with second-order accuracy.
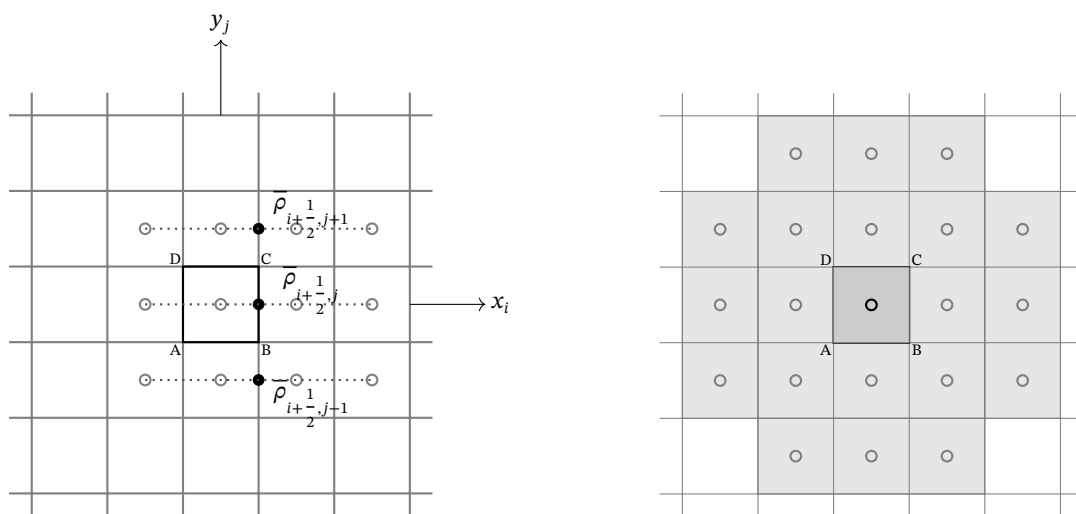
Figure 3.2: *Left* Grid points at cell-centers required to compute $\langle \partial p / \partial x \rangle_{i+1/2,j}$ using (3.18). *Right* A 5×5 block required to compute the flux divergence of cell ABCD.

Boundary conditions

We only explain the discretization at the right and bottom boundaries with Dirichlet and Neumann boundary conditions, respectively. The stencil for the other boundaries can be deduced similarly. Further, we assume the same number of nodes (equal to m) along the x- and y- directions as in the grid definition (3.8). We prefer a ghost cell approach which greatly simplifies the discretization around the faces next to the boundaries. On the right boundary, we use a quartic polynomial extrapolation to update the ghost points $\widetilde{p}$ (see Figure 3.3) at cell centers. In case of Dirichlet boundary conditions, we use:

$$\widetilde{p}_{m+1,j} = \frac{1}{35}\left[128 g_{D_{m+\frac{1}{2},j}} - 140 p_{m,j} + 70 p_{m-1,j} - 28 p_{m-2,j} + 5 p_{m-3,j}\right], \qquad (3.23)$$

where m denotes the index of the cell centers of the control volumes that are located on the boundary. Furthermore, to compute $\overline{\rho}_{m+1/2,j}$ at the face centers of a Dirichlet boundary (labeled as $B1$ in Figure 3.3), we use a cubic polynomial

$$\overline{\rho}_{m+\frac{1}{2},j} = \frac{1}{60}\left[184 g_{D_{m+\frac{1}{2},j}} - 225 p_{m,j} + 50 p_{m-1,j} - 9 p_{m-2,j}\right]. \qquad (3.24)$$

We also consider the Neumann boundary conditions at the bottom boundary. The ghost points are updated using the quartic polynomial given by

$$\widetilde{p}_{i,m+1} = \frac{1}{22}\left[-24 h \cdot g_{N_{i,m+\frac{1}{2}}} + 17 p_{i,m} + 9 p_{i,m-1} - 5 p_{i,m-2} + p_{i,m-3}\right]. \qquad (3.25)$$

At corners (e.g. point $C1$ in Figure 3.3), the ghost points are extrapolated by averaging the values obtained from a quartic polynomial which uses 5 adjacent ghost point values along the two coordinates. The contribution from the x-coordinate reads

$$\widetilde{p}_{m+1,0} = 5\widetilde{p}_{m,0} - 10\widetilde{p}_{m-1,0} + 10\widetilde{p}_{m-2,0} - 5\widetilde{p}_{m-3,0} + \widetilde{p}_{m-4,0}. \qquad (3.26)$$

Quantity $\overline{\rho}_{i,m+1/2} = h \cdot g_{N_{i,m+1/2}}$ at the face-centers of a Neumann boundary. For the corner faces on the boundaries (e.g. the point labeled $B2$ in Figure 3.3), we approximate the transverse Laplacian $\frac{\partial^2 F^{B2}}{\partial x^2}\Big|_{m,1/2}$ and $\frac{\partial^2 p}{\partial y \partial x}\Big|_{i+1/2,j}$ using one-sided finite-differences.
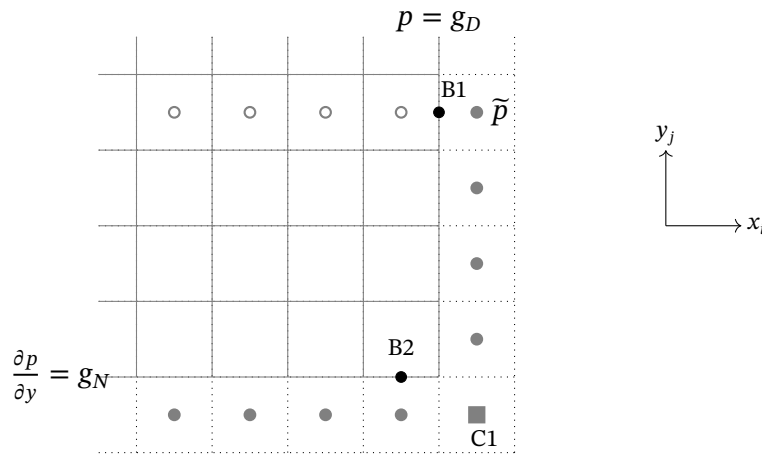
Figure 3.3: Points near or at the boundary which require stencil modification to incorporate boundary conditions and ghost points (denoted by the gray circles and square).

3.4. Full Multigrid algorithm

Now we describe in detail the construction of a geometric FMG solver for the FV discretization developed in the previous section. For the sake of brevity, we will denote the linear system obtained from the second- and fourth-order FV discretizations by

$$\mathcal{L}_h^{(2)} p_h = f_h^{(2)}, \tag{3.27}$$

$$\mathcal{L}_h^{(4)} p_h = f_h^{(4)}, \tag{3.28}$$

respectively. The elements of $\mathcal{L}_h^{(4)}$ are provided in Appendix 3.B. The linear systems (3.27) and (3.28) are solved by the multigrid methods MG2 and MG4, respectively. Further, the full multigrid variants of MG2 and MG4 are denoted by FMG2 and FMG4, respectively.

Typically, a multigrid solver for elliptic problems is easy to construct. A large number of efficient multigrid algorithms can be found in the literature. For a similar problem, the authors in [54, 55] have shown that a multigrid method based on a cell-centered grid using fixed transfer operators that do not depend on matrix coefficients can provide a decent convergence speed, even for highly discontinuous coefficients. They use a second-order FV discretization scheme. In this section we describe a robust and scalable fourth-order accurate multigrid solver.

There are different ways of constructing a multigrid solver for a fourth-order discretization. One way is to follow the standard multigrid approach of using an appropriate smoother and transfer operators. Due to the large size of the fourth-order operator $\mathcal{L}_h^{(4)}$, it becomes difficult to find efficient smoothing schemes compared to the second-order operator $\mathcal{L}_h^{(2)}$. Also, transfer operators will be much more complex, especially, if we are aiming for the independence of coefficient magnitude. A second alternative is to solve the linear system in (3.28) via a *High-Order Defect Correction* (HODC) scheme. This defect correction scheme employs lower-order schemes to obtain higher-order accuracy. Also from a programmer's point of view, it is more convenient to implement the HODC scheme. In our case, we use a second-order multigrid algorithm (MG2) and HODC to construct a fourth-order multigrid solver (MG4). We devote the next section discussing the intricacies.

3.4.1   MG2 cycle

Our MG2 method resembles the cell-centered multigrid (CCMG) algorithm proposed in [54]. Components required to construct a two-grid cycle are discussed below.

Pre- and post-smoothing: There are different possible choices available for the smoothers. In CCMG the ILU smoother works very well but it is quite involved compared to basic iterative smoothers, like Jacobi or Gauss-Seidel smoothers. We will use a Gauss-Seidel Red-Black (GS-RB) smoother in our algorithm.

Defect computation and Restriction: The defect $r_h^k = f_h^{(2)} - \mathcal{L}_h^{(2)} p_h^k$ is computed and is restricted to grid $\mathcal{D}_{2h}$ using a bilinear restriction operator $[R_h^{2h}]$ to obtain r_{2h}^k. The stencil form for this restriction reads:

$$[R_h^{2h}] = \frac{1}{16} \begin{bmatrix} 1 & 1 & & 0 & 0 \\ 1 & 3 & & 2 & 0 \\ & & \star & & \\ 0 & 2 & & 3 & 1 \\ 0 & 0 & & 1 & 1 \end{bmatrix}_h^{2h} .$$

Coarse grid correction: We apply a direct method, e.g. Gaussian elimination, to solve the coarsest grid problem, $\mathcal{L}_{2h}^{(2)} e_{2h}^k = r_{2h}^k$ in a two-grid setting. The prolongation of e_{2h}^k is done using a piece-wise

constant operator $[P_{2h}^h]$ to obtain e_h^k which is added to p_h. The stencil reads:

$$P_{2h}^h = \left. \begin{bmatrix} 1 & & 1 \\ & \star & \\ 1 & & 1 \end{bmatrix} \right]_{2h}^{h}, \tag{3.29}$$

This process is recursively extended to many grid levels. The transfer operator needs to be modified appropriately at the boundaries. The coarse grid operator $\mathcal{L}_{2h}^{(2)}$ is obtained via a direct discretization using an upscaled version of the fine grid coefficient field. The upscaled procedure is outlined in Section 3.7. The coarse grid operator can also be obtained algebraically via the Galerkin coarse grid method where $\mathcal{L}_{2h}^{(2)} = R_h^{2h} \mathcal{L}_h^{(2)} P_{2h}^h$. The Galerkin approach will be discussed in detail in Chapter 4.

3.4.2   MG4 cycle via defect correction

In many scenarios, constructing an iterative solver for lower-order discretizations is comparatively easy. A defect correction strategy [14, 56, 57] can be used to formulate an "outer iteration" where only the right-hand-side of the lower-order linear system is modified on the finest grid using the higher-order linear system for the same problem. As the left-hand-side of the linear system remains the same, a lower-order multigrid cycle can be utilized to perform an "inner iteration". In this case, we consider the second- and fourth-order linear systems defined in (3.27) and (3.28) and the inner iteration uses the MG2 cycle described in the previous subsection. The j-th defect correction iteration is given by

$$\mathcal{L}_h^{(2)} p_h^j = \hat{f}_h, \qquad \text{with} \qquad \hat{f}_h = f_h^{(4)} - \mathcal{L}_h^{(4)} p_h^{j-1} + \mathcal{L}_h^{(2)} p_h^{j-1}. \tag{3.30}$$

We regard this iteration as the MG4 cycle. The new approximation p_h^k is obtained by rearranging:

$$p_h^j = (I - (\mathcal{L}_h^{(2)})^{-1} \mathcal{L}_h^{(4)}) p_h^{j-1} + (\mathcal{L}_h^{(2)})^{-1} f_h^{(4)}. \tag{3.31}$$

Usually, an initial approximation p_h^0 is provided by a solution obtained from a nested FMG solution of second-order accuracy. This scheme converges to the solution of the fourth-order discrete problem if the spectral radius of iteration matrix is strictly less than one, i.e. $\varrho(I - (\mathcal{L}_h^{(2)})^{-1} \mathcal{L}_h^{(4)}) < 1$. This criterion was verified using the Fourier analysis described in [14] for frozen coefficients (for Poisson's equation).

Remark 3.4.1 *It is non-trivial to prove the convergence rate of high-order defect correction scheme for an oscillatory permeability field. We rely on the numerical results to demonstrate the robustness of the solver with respect to covariance parameters λ_c and σ^2.*

3.4.3   FMG structure

Both MG2 and MG4 cycles can be easily incorporated into the FMG-hierarchy to obtain FMG2 and FMG4 solvers, respectively.

Typically, the FMG algorithm can be classified as either "fixed" or "accomodative" [15]. In the former approach, the number of cycles is fixed on all FMG levels beforehand, whereas in the latter approach, the number of cycles is adaptively decided based on some criterion. The accomodative algorithm can be highly effective in optimizing the number of cycles over the FMG levels, especially, for problems involving random coefficient fields. In practice, the error cannot be computed but the optimal switching criteria can be based on residual reduction. The switch from grid $2h \rightarrow h$ is

made when a norm of the residual on grid $2h$ drops below some tolerance ε_{MG}. For example, we can use the constraint

$$\frac{||res^j_{2h}||_2}{||res^0_{2h}||_2} < \varepsilon_{MG}, \tag{3.32}$$

where $||res^0_{2h}||_2$ and $||res^j_{2h}||_2$ are residuals in the 2-norm, computed from the initial solution and after j multigrid cycles (MG2 or MG4) at the $2h$ grid level, respectively. All accommodative algorithms have the disadvantage that they require some extra work of computing residual norms. On the other hand, accommodative algorithms are robust. For a coefficient field with large σ^2 and small λ_c, the linear systems of equations, (3.27) and (3.28), are highly ill-conditioned. In such cases, the residual norm computations are inexpensive relative to other calculations. In the next section, we perform some experiments to study the dependence of the number of cycles required on different grid levels on the regularity of the PDE solution.

Although the accommodative criterion does not always guarantee a reduction of error to the truncation level, it typically can optimize the work on each level. For coupling of the FMG solver with the MLMC estimator, we need to solve the linear system on a sequence of grids. Therefore, the switch to the next grid level is made only after the solution at the current level has reached the discretization accuracy. Currently, there are no theoretical proofs to compute apriori the number of iterations required at FMG levels for problems with random coefficients. Therefore, we use a conservative stopping criterion ($\varepsilon_{MG} < 10^{-6}$) for residuals on these grids in order to ensure that the solution has reached the level of truncation error.

Another key component of an FMG algorithm is the FMG-interpolation. Usually, any high-order scheme gives an accurate solution even on a very coarse grid. Thus, it is important to translate this accuracy to the next finer grid level using an appropriate interpolation scheme for the coarse grid solution. In our algorithm, we use a fourth-order accurate bicubic interpolation (3.33). Due to non-nested nodes in the cell-centered grid hierarchy, we get a relatively large interpolation stencil. This interpolation can be seen as a two-step procedure described in Figure 3.4. First, we generate auxiliary points at the corners and face centers (indicated by the gray circles) of coarse cells using stencil $[\Pi^h_{2h}]^{(1)}$ and $[\Pi^h_{2h}]^{(2)}$, respectively.

$$[\Pi^h_{2h}]^{(1)} = \frac{1}{32}\begin{bmatrix} -1 & 0 & & 0 & -1 \\ 0 & 9 & & 9 & 0 \\ & & \star & & \\ 0 & 9 & & 9 & 0 \\ -1 & 0 & & 0 & -1 \end{bmatrix} \quad \text{and} \quad [\Pi^h_{2h}]^{(2)} = \frac{1}{16}\begin{bmatrix} -1 & 9 & \star & 9 & -1 \end{bmatrix}. \tag{3.33}$$

In step two, we use the auxiliary and coarse grid nodes to generate points that are the fine grid nodes using the same stencil $[\Pi^h_{2h}]^{(1)}$. Near boundaries and corners, appropriate modifications can be used.

3.5. Coupling of MLMC with FMG (FMG-MLMC)

To describe the coupling, we briefly recall the MLMC estimator from Chapter 2. Expectation of a quantity of interest, $Q_{h_L} := Q_{h_L}(p_{h_L})$, can be expressed as

$$\mathbb{E}[Q_{h_L}] = \mathbb{E}[Q_{h_0}] + \sum_{\ell=1}^{L} \mathbb{E}[Q_{h_\ell} - Q_{h_{\ell-1}}], \tag{3.34}$$

and the unbiased MLMC estimator for $\mathbb{E}[Q_{h_L}]$ can be based on standard MC as

$$\mathcal{E}^{ML}_L[Q_{h_L}] = \frac{1}{N_0}\sum_{i=1}^{N_0} Q_{h_0}(\omega_i) + \sum_{\ell=1}^{L}\left(\frac{1}{N_\ell}\sum_{i=1}^{N_\ell}(Q_{h_\ell}(\omega_i) - Q_{h_{\ell-1}}(\omega_i))\right), \tag{3.35}$$
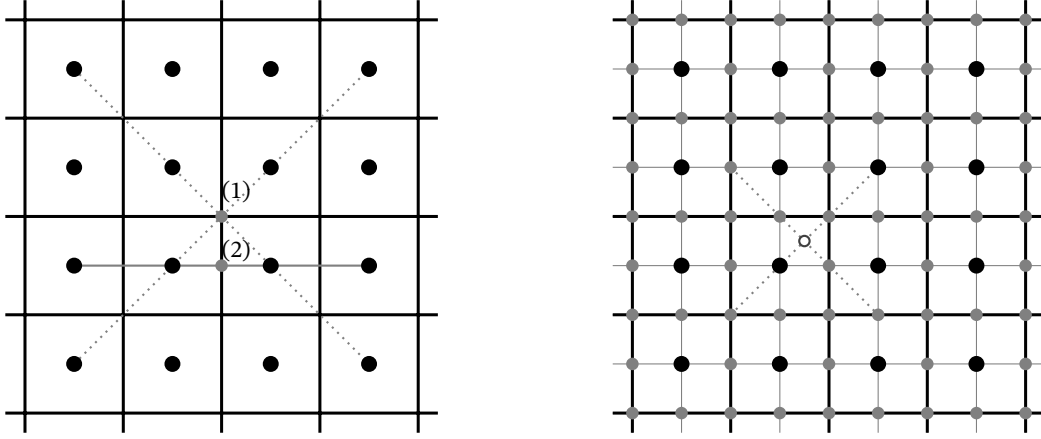
Figure 3.4: Schematic representation of FMG-interpolation procedure *Left:* Generation of auxillary points at corners and face centers of coarse cells *Right:* Interpolation at cell centers of fine grid using coarse grid and auxillary points.

where $\omega_i \in \Omega$. The above estimator is based on a hierachy of cell-centered grids (3.8) with uniform coarsening such that $h_\ell = s^{-1} h_{\ell-1}$ with $s = 2$. We assume that the level-dependent variance $\mathcal{V}_\ell = \mathbb{V}[Q_{h_\ell} - Q_{h_{\ell-1}}] = \mathcal{O}(h_\ell^\beta)$ with $\beta > 0$. The level-dependent samples N_ℓ are determined using the following formula (see Section 2.3.2):

$$N_\ell = \left(\varepsilon^{-2} \sum_{k=0}^{L} \sqrt{\mathcal{V}_k \mathcal{W}_k} \right) \sqrt{\frac{\mathcal{V}_\ell}{\mathcal{W}_\ell}}, \tag{3.36}$$

where ε^2 is the desired MSE and the cost per sample is denoted by $\mathcal{W}_\ell = \mathcal{O}(h_\ell^{-\gamma})$, with $\gamma \geq d$ and d the spatial dimension. The finest grid size is given as $h_L = \mathcal{O}(\varepsilon^{1/\alpha})$, where $\alpha > 0$ is the rate of decay of the discretization bias. More details on the MLMC estimator can be found in Chapter 2.

For every sample $Q_{h_\ell}(\omega_i)$, an expensive linear system of equations obtained from the discretization of the PDE (3.2) needs to be solved. Due to flexibility and ease of use, an optimal "black-box" multigrid solver e.g. AMG1R5 [58] is often preferred for this purpose. One of the disadvantages of such solver is that we do not have access to the solutions on the coarser grid levels. For the MLMC estimator which requires samples at different grid hierarchies, one can benefit from a geometric full multigrid (FMG) solver based on the MLMC grid hierarchy. The efficiency of an FMG solver comes from using an inexpensive-to-compute solution on a coarser grid as a good initial approximation for the solution on the next grid level. There are two-fold benefits of using an FMG solver for an MLMC method. Firstly, these kinds of solvers are asymptotically optimal and can solve the linear system in $\mathcal{O}(h_\ell^{-d})$ operations. Secondly, the solution at each FMG level can be utilized to compute samples of the quantity of interest.

Here we benefit from the fact that the permeability field is generated using the same random sample ω on a sequence of meshes $\{\mathcal{D}_{h_{\ell*}}\}_{\ell*=0}^{\ell}$, $\ell \in \{0, 1, ..., L\}$, and essentially represents the same field, but it is sampled at more nodes as we move to the finer grids. Therefore, the FV approximation of pressure $p_{h_{\ell*}}$ on the level ℓ^* can be utilized to accelerate the solve on the next finer level $\ell^* + 1$ and at the same time can be used to compute $Q_{h_{\ell*}}$.

We use the notation FMG^ℓ to denote an FMG solver with grid levels $\{h_{\ell*}\}_{\ell*=0}^{\ell}$. We now demonstrate how to compute a sample of the quantity of interest at all MLMC levels using the FMG^L solver. First the permeability field is generated at all levels using the same random sample ω_i, $i = 1, 2, ..., N_L$. Using this field, we formulate the linear system at each grid level of FMG^L, i.e.

$$\mathcal{L}_{h_\ell}(\omega_i) p_{h_\ell} = f_{h_\ell} \quad \text{for} \quad \ell \in \{0, 1, ..., L\}. \tag{3.37}$$
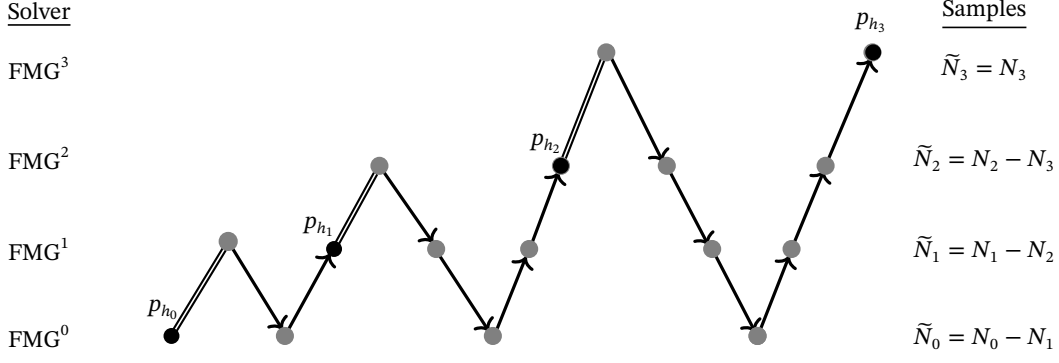
Figure 3.5: Schematic representation of a 4-level FMG-MLMC method. The black dots represent the converged solution points and the double lines represent the FMG-interpolation Π_{2h}^h. FMG solvers are listed on left and the number of samples to be computed from these solvers is listed on the right.

We begin by solving the coarsest grid problem, $\mathcal{L}_{h_0}(\omega_i)p_{h_0} = f_{h_0}$ upto the discretization accuracy and computing the sample $Q_{h_0}(\omega_i)$. Then, the interpolated version of this solution $\Pi_{2h}^h p_{h_0}$, where Π_{2h}^h is the FMG-interpolation operator, is utilized as an initial approximation for the problem on the next level $\ell = 1$. Again, after a few multigrid cycles we obtain the solution p_{h_1} and subsequently compute $Q_{h_1}(\omega_i)$. This is done recursively till the finest level L. This way, based on one FMG solve, we get one sample of the quantity of interest at each grid level i.e. $Q_{h_0}(\omega_i), (Q_{h_1}(\omega_i) - Q_{h_0}(\omega_i)), ..., (Q_{h_\ell}(\omega_i) - Q_{h_{\ell-1}}(\omega_i)), ..., (Q_{h_L}(\omega_i) - Q_{h_{L-1}}(\omega_i))$. Repeating this process for N_L independent realizations of the random field, we get N_L samples *at all MLMC levels.* Now, for the next coarser MLMC level $L-1$, we just need to compute $\widetilde{N}_{L-1} = (N_{L-1} - N_L)$ extra samples using the solver FMG^{L-1}. These samples are computed in similar fashion as above. We note that, for any level $0 \leq \ell < L$, we already have $N_{\ell+1}$ samples from higher levels $\ell + 1, ..., L$. Therefore, the remaining samples $\widetilde{N}_\ell$ to be computed at level ℓ are given by

$$\widetilde{N}_\ell = N_\ell - N_{\ell+1}. \tag{3.38}$$

Note that on the finest level MLMC level L, we have $\widetilde{N}_L = N_L$. A straightforward computation using (3.36) shows that for the levels $0 < \ell < L$, the number of samples grows as:

$$N_\ell = N_{\ell+1} s^{\frac{(\beta+\gamma)}{2}}, \tag{3.39}$$

where s is the grid scaling factor defined earlier. From (3.38) and (3.39), we get the reduction in the number of samples as:

$$\frac{\widetilde{N}_\ell}{N_\ell} = 1 - s^{-\frac{(\beta+\gamma)}{2}}. \tag{3.40}$$

This shows that the reduction in the number of samples is more pronounced when the rates β and γ are small. In other words, the "recycling" of the coarse grid samples is more effective when N_ℓ decays slowly with level. Figure 3.5 represents a 4-level FMG-MLMC method using a simplified FMG cycle. The computational cost of the FMG-MLMC estimator is calculated as $\sum_{\ell=0}^L \widetilde{N}_\ell \mathcal{W}_\ell$. For comparison, the cost of the standard MLMC estimator based on the "black-box" approach will be $N_0 \mathcal{W}_0 + \sum_{\ell=1}^L N_\ell(\mathcal{W}_\ell + \mathcal{W}_{\ell-1})$. This shows that the FMG approach will result in higher computational gain.

In practice, the FMG-MLMC algorithm can be implemented by slightly modifying the MLMC algorithm described in Section 2.3.2, see also [10, 28]. The implementation details are provided in Algorithm 2.

Algorithm 2 FMG-MLMC algorithm

1: Start with $L' = 1$ and run some initial number of FMG solves at level L'.
2: Estimate the sample variance $\mathcal{V}_\ell, \ell = 0, 1, ..., L'$ by using the available number of samples.
3: Using (3.36) compute the optimal number of samples, N_ℓ for all $\ell = 0, 1, ..., L'$ and then compute $\widetilde{N}_\ell = N_\ell - N_{\ell-1}$.
4: Evaluate the extra $\widetilde{N}_\ell$ samples by using the FMG$^\ell$ solver and update $\mathcal{V}_\ell$ for all levels.
5: Test for convergence using the criterion $|\mathbb{E}[Q_{h_L} - Q_{h_{L-1}}]| < (s^\alpha - 1)\frac{\varepsilon}{\sqrt{2}}$, if converged, set $L = L'$.
6: If not converged, set $L' = L' + 1$ and go back to step 2.

When using this procedure it is important to ensure that the telescopic identity (3.35) is not violated. For this, we need to confirm that the definitions of Q_{h_ℓ} when computing $\mathbb{E}[Q_{h_\ell} - Q_{h_{\ell-1}}]$ and $\mathbb{E}[Q_{h_{\ell+1}} - Q_{h_\ell}]$ have the same expectation, i.e.

$$\mathbb{E}[Q_{h_\ell}]^{(coarse)} = \mathbb{E}[Q_{h_\ell}]^{(fine)} \quad \text{for} \quad \ell \in \{0, 1, ..., L-1\}. \tag{3.41}$$

The standard FMG approach where the coarse grid problem is formulated using Galerkin coarsening or a direct discretization of the problem with a restricted coefficient field from fine grid may lead to some additional bias which may decay at a slower rate than the FV error itself. Therefore, to avoid this, we recommend generating the coefficient field on coarse grids using the same random vector ω_i. We will present one way of doing this in Section 3.7.

The coarsest level in the FMG solver is decided on the basis of certain stability criteria. In general, the coarsest level should be able to provide a minimum level of resolution to the problem such that it serves as a meaningful initial guess for the next level in the FMG hierarchy. In particular, for the groundwater flow problem the coarsest mesh size should be of the order of the correlation length, i.e., $h_0 = \mathcal{O}(\lambda_c)$.

Remark 3.5.1 *We point out that due to sample reuse in the FMG-MLMC estimator, samples on different levels are correlated. Therefore, when estimating the sampling error for this estimator, the covariance between different levels should also be incorporated. Obtaining an analytic bound for the FMG-MLMC estimator with correlated samples is not straightforward. In [59] the authors proposed a valid way of estimating sampling error by taking multiple independent estimators. The paper uses multiple QMC estimators to approximate variance, however, this approach also works for the standard MC.*

3.6. Sampling Gaussian random fields

Several techniques exist to generate samples of Gaussian random fields such as the Cholesky decomposition, the truncated Karhunen-Loéve (KL) expansion [5] or using spectral generators (circulant embedding techniques) [60–62]. The realization of Z can be based on the Karhunen-Loéve (KL) decomposition:

$$Z(\mathbf{x}, \omega) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} \Psi_j(\mathbf{x}) \xi_j, \qquad \xi_j \sim \mathcal{N}(0, 1). \tag{3.42}$$

Here, λ_j and Ψ_j are eigenvalues and eigenfunctions of the covariance kernel $C_\Phi(\mathbf{x_1}, \mathbf{x_2})$, obtained from the solution of the Fredholm integral,

$$\int_{\mathcal{D}} C_\Phi(\mathbf{x_1}, \mathbf{x_2}) \Psi(\mathbf{x_1}) d\mathbf{x_1} = \lambda \Psi(\mathbf{x_2}). \tag{3.43}$$

The sum in (3.42) represents an infinite dimensional uncertain field with diminishing contributions of the eigenmodes. The sum is truncated after a finite number of terms, M_{KL}, which is usually

decided by balancing the KL truncation error with other sources of error, such as discretization or sampling errors. For Gaussian processes with small correlation lengths and large variances, typically a large number of terms is needed to include all important eigenmodes [5]. The evaluation of eigenmodes in the KL expansion is expensive as it requires solving the integral equation (3.43) for each mode.

When using a stationary covariance model, a fast sampling of random fields can be achieved via a spectral generator which employs the discrete FFT (Fast Fourier Transform). Contrary to the KL expansion, spectral generators provide an exact representation of random fields on the sampling mesh. There are several ways of constructing a spectral generator, we use the Fast Fourier Transform moving average (FFT-MA) technique from [60] to decompose the covariance matrix $C_\Phi(r)$. Although this sampling method is similar to the Cholesky factorization technique, the key idea is to make the computational domain periodic. Thus, the resulting covariance operator is also periodic, which can now be decomposed as a convolutional product. The samples of random fields are computed using cheaper vector-vector products compared to the expensive matrix-vector operation required when using the Cholesky factorization. As a periodic covariance function sampled on a uniform grid results in a *circulant* covariance matrix, these methods are sometimes also referred to as the circulant embedding technique. In the following, we provide a brief description of the FFT-MA method from [60].

Let us denote by $\mathbf{z}_h = Z(\mathbf{x}_h, \omega) \in \mathbb{R}^M$ the correlated Gaussian random vector sampled on mesh $\mathcal{D}_h$ with $M = m \times m$ as the total number of grid points (cf. (3.8)). Using a Cholesky decomposition of the covariance matrix $\mathbf{C}_\Phi^h \in \mathbb{R}^{M \times M}$, one can obtain $\mathbf{z}_h$:

$$\mathbf{C}_\Phi^h = \mathbf{U}_h \mathbf{U}_h^T \quad \text{and use} \quad \mathbf{z}_h = \mathbf{U}_h \mathbf{y}_h, \tag{3.44}$$

where $\mathbf{U}_h \in \mathbb{R}^{M \times M}$ is an upper triangular matrix with positive diagonal entries and $\mathbf{y}_h \in \mathbb{R}^M$ is a vector of i.i.d. samples from the standard normal distribution. The FFT-MA method is based on a decomposition of the covariance function $C_\Phi(r)$ (recall (3.7)) as a convolutional product of some function $S_\Phi(r)$ and its transpose $S'_\Phi(r) := S_\Phi(-r)$. In a discrete setting, we can express this decomposition as

$$\mathbf{c}_h = \mathbf{s}_h * \mathbf{s}'_h, \tag{3.45}$$

where $*$ denotes the convolution product and the vectors $\mathbf{c}_h, \mathbf{s}_h$ are vectors obtained by evaluating $C_\Phi(r)$ and $S_\Phi(r)$, respectively, at grid points of the mesh $\mathcal{D}_h$. A correlated random vector $\mathbf{z}_h$ can now be synthesized by using the convolution product,

$$\mathbf{z}_h = \mathbf{s}_h * \mathbf{y}_h. \tag{3.46}$$

The FFT-MA method performs the above computations in the frequency domain. As the FFT requires periodicity, first the vector $\mathbf{c}_h$ is transformed into a periodic signal, which is also real, positive and symmetric. More details on the practical aspects of this transformation, see for e.g. [38, 61, 62]. Also the resulting vector $\mathbf{s}_h$ is real, positive and symmetric and $\mathbf{s}_h = \mathbf{s}'_h$. As the convolution product in the spatial domain is equivalent to the component-wise product in the frequency domain, we can rewrite (3.45) as

$$\overline{\mathcal{F}}(\mathbf{c}_h) = \overline{\mathcal{F}}(\mathbf{s}_h) \cdot \overline{\mathcal{F}}(\mathbf{s}_h) \implies \overline{\mathcal{F}}(\mathbf{s}_h) = \sqrt{\overline{\mathcal{F}}(\mathbf{c}_h)}, \tag{3.47}$$

where $\overline{\mathcal{F}}$ denotes the discrete FFT and $\cdot$ denotes component-wise multiplication. It is pointed out that the component-wise square-root operation does not pose any problems as the power spectrum $\overline{\mathcal{F}}(\mathbf{c}_h)$ is real, positive. Next, we express the convolution product in (3.46) as a vector-vector product in the frequency domain as

$$\overline{\mathcal{F}}(\mathbf{z}_h) = \overline{\mathcal{F}}(\mathbf{s}_h * \mathbf{y}_h) = \overline{\mathcal{F}}(\mathbf{s}_h) \cdot \overline{\mathcal{F}}(\mathbf{y}_h). \tag{3.48}$$

In the final step, an inverse fast Fourier transform $\overline{\mathcal{F}}^{-1}$ is applied to obtain the samples for the Gaussian random fields

$$\mathbf{z}_h = \overline{\mathcal{F}}^{-1}(\overline{\mathcal{F}}(\mathbf{s}_h) \cdot \overline{\mathcal{F}}(\mathbf{y}_h)). \tag{3.49}$$

Due to the periodicity in the covariance vector $\mathbf{c}_h$, the resulting random field $\mathbf{z}_h$ is also periodic. Thus, we only retain the part of the vector that corresponds to the physical domain. We also remark that it takes two FFT evaluations to obtain one sample of $\mathbf{z}_h$ (ignoring the FFT operation in (3.47) that is performed just once). Therefore, in terms of the number of floating point operations, the sampling cost is significantly smaller compared to one multigrid solve for the mesh sizes considered.

3.7. Upscaling Gaussian random fields

While estimating the correction term $\mathbb{E}[Q_{h_\ell} - Q_{h_{\ell-1}}]$ in the telescopic sum (3.34), the approximations $Q_{h_\ell}(\omega_i)$ and $Q_{h_{\ell-1}}(\omega_i)$ need to be positively correlated such that the variance $\mathbb{V}[Q_{h_\ell} - Q_{h_{\ell-1}}]$ is small. This is typically achieved by first sampling the fine grid Gaussian random field, $Z(\mathbf{x}_{h_\ell}, \omega_i)$ to compute $Q_{h_\ell}(\omega_i)$ and using an upscaled version, $Z(\mathbf{x}_{h_{\ell-1}}, \omega_i)$ for $Q_{h_{\ell-1}}(\omega_i)$. Many of the upscaling algorithms based on homogenization techniques in the context of deterministic PDEs, such as [63, 64], cannot be directly applied to obtain the upscaled permeability. This is because these homogenization procedures may result in a modified covariance structure on the coarser levels. When using the KL expansion an upscaled random field can be obtained by using the same Gaussian vectors $\{\xi_j\}_{j=1}^{M_{KL}}$ in the expansion on both levels as:

$$Z(\mathbf{x}_{h_\ell}, \omega) = \sum_{j=1}^{M_{KL}} \sqrt{\lambda_j} \Psi_j(\mathbf{x}_{h_\ell}) \xi_j, \tag{3.50}$$

$$Z(\mathbf{x}_{h_{\ell-1}}, \omega) = \sum_{j=1}^{M_{KL}} \sqrt{\lambda_j} \Psi_j(\mathbf{x}_{h_{\ell-1}}) \xi_j, \tag{3.51}$$

When using the spectral generator, the coarse grid random field can be obtained by using the *covariance upscaling* [32] that employs the spectral generator on two consecutive grids using the same normally distributed vector $\mathbf{y}_{h_\ell}$ (from Eq. (3.44)). As in the case of the FFT-MA algorithm, the vector $\mathbf{y}_{h_\ell}$ is associated with respective grid points, a coarser realization $\mathbf{z}_{h_{\ell-1}}$ of the fine grid Gaussian random field $\mathbf{z}_{h_\ell}$ can be obtained by using multi-dimensional averaging of the vector $\mathbf{y}_{h_\ell}$. For instance, in two dimensions for the cell-centered mesh,

$$\mathbf{y}_{h_{\ell-1}}^{i,j} = \frac{1}{2}(\mathbf{y}_{h_\ell}^{2i-1,2j-1} + \mathbf{y}_{h_\ell}^{2i-1,2j} + \mathbf{y}_{h_\ell}^{2i,2j-1} + \mathbf{y}_{h_\ell}^{2i,2j}), \tag{3.52}$$

where (i, j) is the cell index for the mesh $\mathcal{D}_{h_{\ell-1}}$. The scaling by a factor 2 is needed to obtain a standard normal distribution for the averaged quantity $\mathbf{y}_{h_{\ell-1}}^{i,j}$. The coarse random field can now be simply assembled as

$$\mathbf{z}_{h_{\ell-1}} = \overline{\mathcal{F}}^{-1}(\overline{\mathcal{F}}(\mathbf{s}_{h_{\ell-1}}) \cdot \overline{\mathcal{F}}(\mathbf{y}_{h_{\ell-1}})). \tag{3.53}$$

This process can be recursively applied to generate upscaled random fields for all coarser levels of the FMG solver. As the averaging in (3.52) smooths out high frequencies, the upscaled version $\mathbf{z}_{h_{\ell-1}}$ will also be smoother compared to $\mathbf{z}_{h_\ell}$.

3.8. Numerical results

In this section we examine in detail the performance of various components described in the preceding sections. We consider PDE (3.2) on domain $\mathcal{D} \in (0, 1)^2$ with a mixed Dirichlet-Neumann boundary conditions,

$$g_D(0, y) = 1, \quad g_D(1, y) = 0, \quad \text{and} \tag{3.54}$$

$$g_N(x, 0) = 0, \quad g_N(x, 1) = 0, \quad \text{for} \quad x, y \in (0, 1), \tag{3.55}$$

respectively. For all tests we consider $f = 0$ so that the regularity of p only depends on the Matérn parameters. All numerical schemes are implemented in MATLAB and results are generated on a common workstation.

3.8.1 Multigrid convergence

Our first task is to illustrate the convergence rates of the 2D multigrid solver for the second- and fourth-order FV discretizations, as described in the preceding sections. A theoretical convergence analysis for the exponential covariance ($\nu_c = 1/2$) has been conducted in [65] for two-level multigrid along with different choices of smoothers. Here, we consider challenging cases with higher variances and small correlation lengths. For this we choose 6 different combinations of the Matérn parameters (cf. Table 3.1) with increasing order of complexity in terms of solvability of the linear system. Also, we show one random realization of the Gaussian random field generated using each of the six Matérn parameters in Figure 3.6. These examples clearly show that for the considered Matérn parameters, the simulated permeability fields can exhibit large fluctuations and a variation of more that 10 orders of magnitude can be encountered for the cases with $\sigma_c^2 = 3$.

Table 3.1: Different combinations of the Matérn parameters $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$ with increasing complexity from left to right.

Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6
$(2.5, 0.3, 1)$	$(1.5, 0.3, 1)$	$(0.5, 0.3, 1)$	$(2.5, 0.1, 3)$	$(1.5, 0.1, 3)$	$(0.5, 0.1, 3)$

For a fixed Φ and h we generate 100 samples of the random field. Then, for each sample we run the multigrid as well as the FMG cycles. The reduction factor of a multigrid cycle for the i-th realization of the random field is defined as

$$\rho_i := \left\{ \frac{||res_h^{\kappa_i}||_2}{||res_h^0||_2} \right\}^{1/\kappa_i}, \qquad \text{for} \qquad i = 1, 2, ..., 100, \tag{3.56}$$

with $||res_h^0||_2$ the 2-norm of the residual from a zero initial guess and $||res_h^{\kappa_i}||_2$ the residual after κ_i multigrid iterations required to achieve the reduction, $||res_h^{\kappa_i}||_2/||res_h^0||_2 < 10^{-6}$. Finally, the average reduction factor is computed as

$$\langle \rho \rangle = \frac{1}{100} \sum_{i=1}^{100} \rho_i. \tag{3.57}$$

In all the test cases, we only consider the multigrid V(1,1)-cycle. Further, we fix the coarsest grid based on the correlation length with $h_0 = 1/16$ for $\lambda_c = 0.3, \sigma_c^2 = 1$ and $h_0 = 1/32$ for $\lambda_c = 0.1, \sigma_c^2 = 3$. First, we consider the multigrid cycle MG2 for a second-order accurate solution and the corresponding full multigrid cycle FMG2. The reduction rate of a multigrid V-cycle is governed by the number of multigrid levels as well as by how well the random field is resolved at the multigrid levels. Figure 3.7 (*left*) presents the average reduction rates, $\langle \rho \rangle$, with different mesh widths. Additionally, Table 3.2 provides the numerical value along with the observed standard deviation, σ, for the MG2 algorithm. For each parameter set, the average reduction factor is roughly the same for different grid sizes with slight improvement as we move to a finer grid. We also see
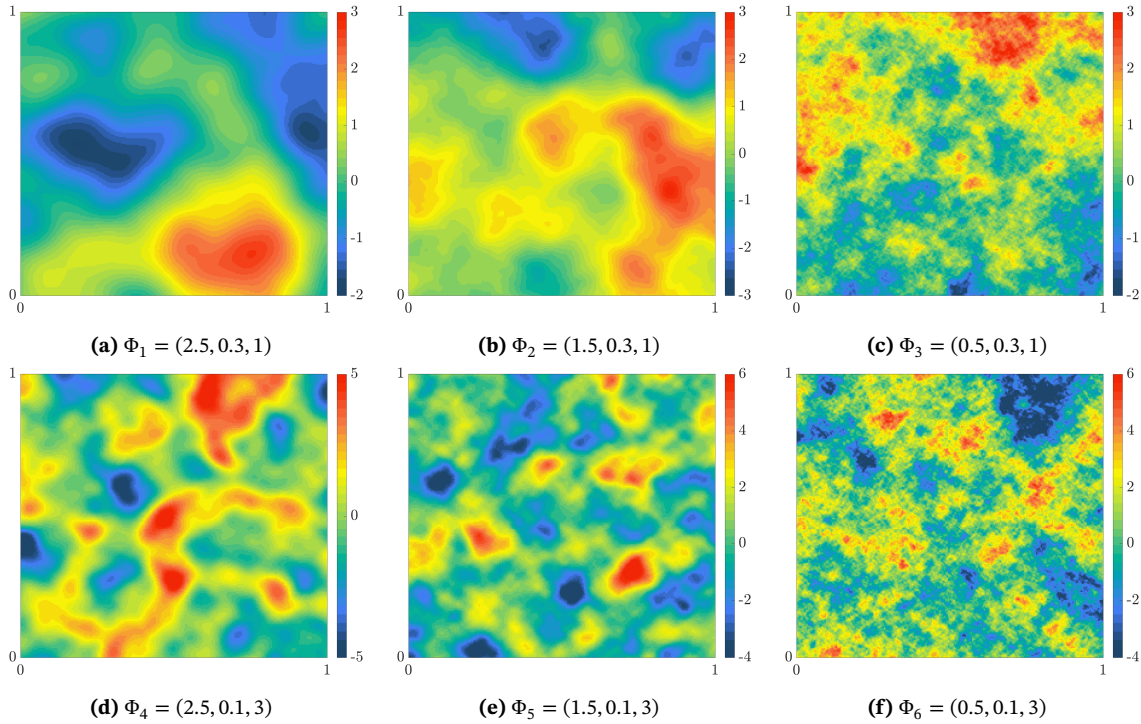
(a) $\Phi_1 = (2.5, 0.3, 1)$ **(b)** $\Phi_2 = (1.5, 0.3, 1)$ **(c)** $\Phi_3 = (0.5, 0.3, 1)$

(d) $\Phi_4 = (2.5, 0.1, 3)$ **(e)** $\Phi_5 = (1.5, 0.1, 3)$ **(f)** $\Phi_6 = (0.5, 0.1, 3)$

Figure 3.6: Examples of Gaussian random field ($\log k$) generated using parameter sets $\Phi_1 - \Phi_6$. The colour bars indicate the variation in the order of magnitude.

that the deterioration in the average reduction factor is more pronounced with increasing variance compared to the decreasing smoothness and correlation length. In order to measure the efficiency of the FMG2 cycle, we run a single cycle and check for the residual reduction. If not converged then extra V-cycles are run which is denoted by κ^{FMG2}. Table 3.3 provides the bound for the average numbers of MG2[FMG2] iterations, $\langle \kappa^{MG2} \rangle [\langle \kappa^{FMG2} \rangle]$, required to reduce the residual by a factor of 10^{-6}. Also, note that the FMG2 method improves as we move to the finer grids for all cases.

Similarly, we investigate the reduction rates of the multigrid cycle MG4 for a fourth-order accurate solution along with its full multigrid variant FMG4, based on the defect correction strategy. We follow a similar procedure as above to measure the performance of these solvers. Also, we use the fourth-order solver only for the cases with $\nu_c > 1$ as we do not expect any improvement for the rough cases (Φ_3, Φ_6) over the second-order discretization. Table 3.4 reports the average reduction rates to achieve the stopping criteria along with the standard deviations. From Figure 3.7 (*right*), we observe that the average reduction rate for MG4 cycle stabilizes around 0.45. The better performance for cases Φ_4, Φ_5 can be attributed to the use of $h_0 = 1/32$ compared to Φ_1, Φ_2 which use $h_0 = 1/16$. Table 3.5 provides the number of defect correction steps $\langle \kappa^{MG4} \rangle$ for the MG4 method. Using an initial guess from an FMG2 cycle instead of using a zero initial guess can reduce the number of iterations. Next, the quantity $\langle \kappa^{FMG4} \rangle$ shows the dependence of the defect correction steps combined with the FMG4 solver on the regularity of the solution. As expected, more iterations are required when the regularity decreases. Interestingly, the FMG4 solver is very efficient as it converges to the stopping criterion in one cycle on finer grids, which is a huge improvement when compared to the number of MG4 iterations that falls around 20. Note that for finer grids such as $1/512$, the residual reduction by a factor of 10^{-6} may not be enough to reach the discretization accuracy and should be further lowered.

A few remarks are in order. Improvement in the average reduction rate was observed with the F(1,1)- and W(1,1)-cycles but V(1,1)-cycle was the fastest to reach the stopping criteria. The combination $\sigma_c^2 = 3$ and $\lambda_c = 0.1$, using $h_0 = 1/16$, resulted in a very slow convergence and in

divergence for some cases. Further, on coarser grids (till 64x64), the performance of the sparse direct solver in MATLAB is superior in terms of CPU times compared to the MG4/FMG4 solvers. This is due to the dominating setup cost compared to the MG4 iterations itself.
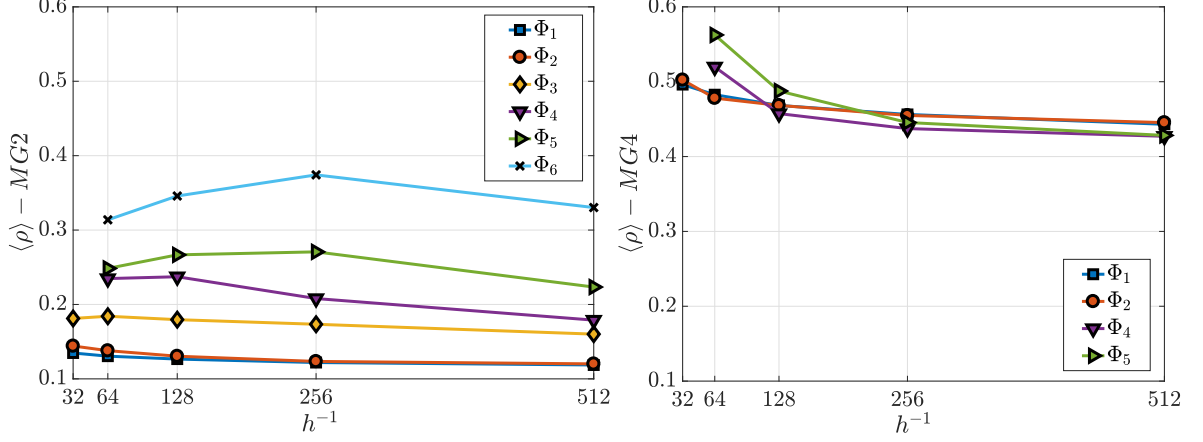

Figure 3.7: Average reduction factor for MG2- and MG4-cycles for different h.

Table 3.2: Average reduction factor [standard deviation], $\langle \rho \rangle [\sigma]$, for different grid sizes for MG2 cycles.

Φ	$h = 1/32$	$h = 1/64$	$h = 1/128$	$h = 1/256$	$h = 1/512$
Φ_1	0.134[0.013]	0.130[0.010]	0.126[0.008]	0.122[0.006]	0.118[0.003]
Φ_2	0.144[0.031]	0.137[0.017]	0.130[0.010]	0.123[0.005]	0.120[0.007]
Φ_3	0.181[0.025]	0.184[0.025]	0.179[0.028]	0.173[0.022]	0.160[0.017]
Φ_4	-	0.234[0.078]	0.237[0.130]	0.208[0.106]	0.178[0.075]
Φ_5	-	0.248[0.087]	0.266[0.125]	0.270[0.169]	0.223[0.123]
Φ_6	-	0.313[0.061]	0.345[0.097]	0.374[0.149]	0.330[0.086]

Table 3.3: Average number of iterations for MG2 [FMG2], $\langle \kappa^{MG2} \rangle [\langle \kappa^{FMG2} \rangle]$, for different grid sizes.

Φ	$h = 1/32$	$h = 1/64$	$h = 1/128$	$h = 1/256$	$h = 1/512$
Φ_1	8[3]	8[2]	8[2]	8[1]	7[1]
Φ_2	8[4]	8[3]	8[2]	7[2]	8[1]
Φ_3	9[5]	9[4]	9[4]	9[3]	8[2]
Φ_4	-	11[5]	12[5]	10[2]	9[2]
Φ_5	-	11[5]	13[5]	14[5]	11[3]
Φ_6	-	13[7]	15[7]	18[8]	14[4]

3.8.2 Convergence of the FMG-MLMC method

In this section, we test the convergence of the FMG-MLMC methods outlined in the previous sections. We denote by FMG2-MLMC and FMG4-MLMC, the multilevel estimators obtained from second- and fourth-order discretization, respectively. Further, we denote by MC2 and MC4 the single-level Monte Carlo estimators using the second- and fourth-order discretizations, respectively.

Figure 3.8 shows the average CPU time required to solve the second- and fourth-order linear systems in 2D using the full multigrid solver for one random sample. The time required to generate

Table 3.4: Average reduction factor [standard deviation], $\langle\rho\rangle[\sigma]$, for different grid sizes for MG4 cycles.

Φ	$h = 1/32$	$h = 1/64$	$h = 1/128$	$h = 1/256$	$h = 1/512$
Φ_1	0.496[0.038]	0.483[0.026]	0.468[0.021]	0.456[0.017]	0.443[0.019]
Φ_2	0.502[0.041]	0.478[0.030]	0.468[0.025]	0.455[0.021]	0.445[0.021]
Φ_4	-	0.520[0.041]	0.457[0.030]	0.437[0.025]	0.427[0.021]
Φ_5	-	0.562[0.082]	0.488[0.070]	0.446[0.047]	0.428[0.041]

Table 3.5: Average number of iterations for MG4 [FMG4], $\langle\kappa^{MG4}\rangle[\langle\kappa^{FMG4}\rangle]$, for different grid sizes.

Φ	$h = 1/32$	$h = 1/64$	$h = 1/128$	$h = 1/256$	$h = 1/512$
Φ_1	21[2]	20[2]	19[1]	19[1]	18[1]
Φ_2	21[3]	20[2]	19[1]	19[1]	18[1]
Φ_4	-	23[3]	19[2]	18[1]	18[2]
Φ_5	-	26[5]	21[2]	18[2]	17[2]

one random sample by the circulant embedding technique is also provided. We note that the cost of solving the linear system dominates for the given range of h_ℓ. Therefore, in all our numerical results we will assume that the cost to compute one sample grows with the rate $\gamma \approx d = 2$. Since our code is not optimized, we will use a standardized cost model where we set the cost to compute one sample, $\mathcal{W}_\ell = h_\ell^{-d}$, in order to compare the costs of different estimators rather than using the CPU times.
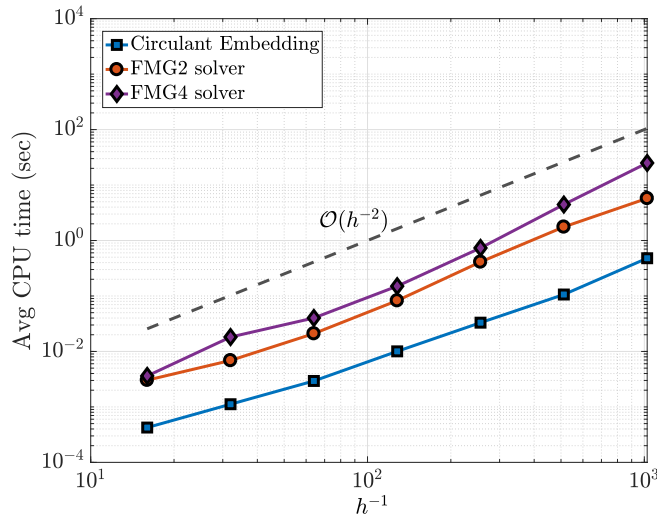
Figure 3.8: Average CPU times (in sec) for the solution of second- and fourth-order discretizations along with the time to generate one sample of random field via the circulant embedding method for 2D problem with $\Phi = (1.5, 0.1, 1)$.

We will now compute the expected values of the different quantities of interest using the FMG-MLMC estimators. We refer readers to [39, 41, 66] for the convergence proofs for different linear functionals of the solution p of the Darcy flow equation. The first output quantity of interest is the horizontal flux at the center of domain $\mathbf{x}^* = (1/2, 1/2)^T$, i.e.

$$Q(p) = -k(\mathbf{x}, \omega) \left. \frac{\partial p(\mathbf{x}, \omega)}{\partial x} \right|_{(\mathbf{x}^*)}, \tag{3.58}$$

where the partial derivative $\frac{\partial p(\mathbf{x}, \omega)}{\partial x}$ is computed using Taylor's expansion with $\mathcal{O}(h^2)$ accuracy. In
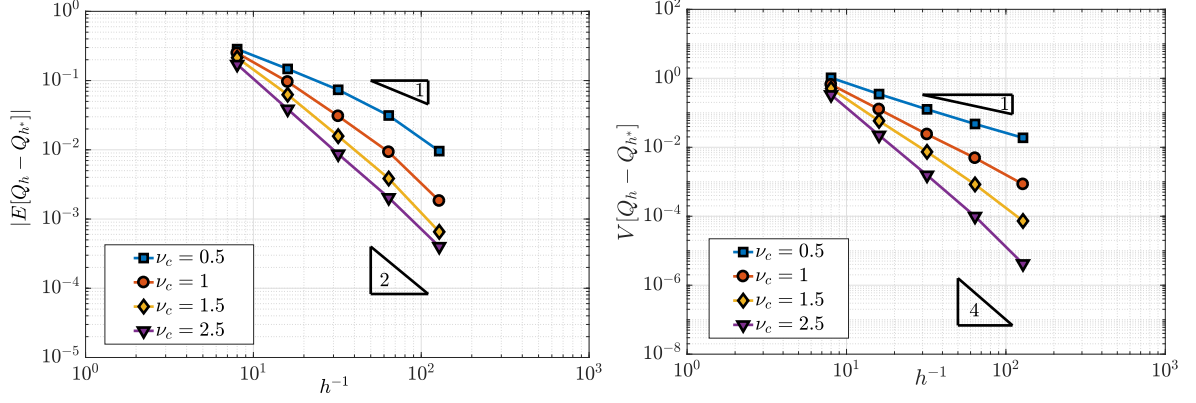
Figure 3.9: Parameter α *(left)* and β *(right)* observed for horizontal flux $-k\left.\frac{\partial p}{\partial x}\right|_{(\mathbf{x}^*)}$ where $\mathbf{x}^* = (1/2, 1/2)^T$ from the FMG2-MLMC scheme with $\lambda_c = 0.1$, $\sigma_c^2 = 1$ for different smoothness parameters ν_c with the reference grid $h^* = 1/256$. The integer in the triangles illustrates the convergence rates $w.r.t$ to h.

Figure 3.9, we show the convergence rate of the FMG2-MLMC method with $\lambda_c = 0.1$, $\sigma_c^2 = 1$ and smoothness parameter $\nu_c \in \{0.5, 1, 1.5, 2.5\}$. For each ν_c, we compute the expectation $|\mathbb{E}[Q_{h_\ell} - Q_{h^*}]|$ and the variance $\mathbb{V}[Q_{h_\ell} - Q_{h^*}]$ from 10,000 independent realizations of the random field using the FMG2 solver with the finest grid $h^* = 1/256$. The purpose here is to show that the approximation at each FMG level has reached the discretization accuracy and can be used as valid sample for the MLMC estimator. Also, we see that the rate of convergence α (left) of the expectation, $|\mathbb{E}[Q_{h_\ell} - Q_{h^*}]|$, improves with increasing smoothness but stalls after $\nu_c > 1$. The convergence rate β (right) of the variance, $\mathbb{V}[Q_{h_\ell} - Q_{h^*}]$, does however improve with smoothness until β reaches the value 4 which follows from the inequality $\mathbb{V}[Q] \leq \mathbb{E}[Q^2]$. This verifies the inability of second-order schemes to capture additional regularity in the solution obtained when $\nu_c > 1$.

In Figure 3.10, we use the FMG4-MLMC method for the same quantity of interest and compare with the FMG2-MLMC method for $\nu_c = 1.5$ (top) and $\nu_c = 2.5$ (bottom). Again, the partial derivative $\partial p(\mathbf{x}, \omega)/\partial x$ is computed using Taylor's expansion with $\mathcal{O}(h^4)$ accuracy. To differentiate between the two methods, we denote the convergence rates by $\alpha^{(2)}$, $\beta^{(2)}$ and $\alpha^{(4)}$, $\beta^{(4)}$ for second- and fourth-order, respectively. The empirical values of these parameters obtained from regression are listed on the top of each figure. We observe that $\alpha^{(4)} > \alpha^{(2)}$ and $\beta^{(4)} > \beta^{(2)}$ for both smoothness values and the difference is more pronounced with increasing smoothness of the random field. Recall that a higher value of α implies fewer MLMC levels and a higher β corresponds to fewer MLMC samples on those levels.

In Figure 3.11, we compare the standardized cost to compute the expected values of the horizontal flux at the center with different tolerances ε. We use the coarsest grid $h_0 = 1/8$ for both FMG2-MLMC and FMG4-MLMC estimators. These results are produced using Algorithm 2 provided in Section 3.5. The estimators FMG2-MLMC and FMG4-MLMC converge with the rate $\mathcal{O}(\varepsilon^{-2})$ consistent with the MLMC theorem ($\beta > \gamma = 2$). The cost of the FMG4-MLMC estimator is however significantly lower than the other estimators which can be attributed to higher α and β. The MC2 method converges with the expected rate $\mathcal{O}(\varepsilon^{-3})$, whereas MC4 converges to $\mathcal{O}(\varepsilon^{-2.8})$ *(left)* and improves slightly with $\mathcal{O}(\varepsilon^{-2.55})$ *(right)* for the smoother test case. Next, we consider the mean of the solution p in $\mathcal{D}$,

$$Q(p) = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} p(\mathbf{x}, \omega) d\mathbf{x}. \tag{3.59}$$

The above integral is approximated using a fourth-order quadrature rule. In Figure 3.12, we compare the convergence rates of the two methods for smoothness parameter $\nu_c = \{1.5, 2.5\}$. The top two plots show the convergence rates for the relatively "easier" Matérn parameter set $\Phi = (2.5, 0.3, 1)$
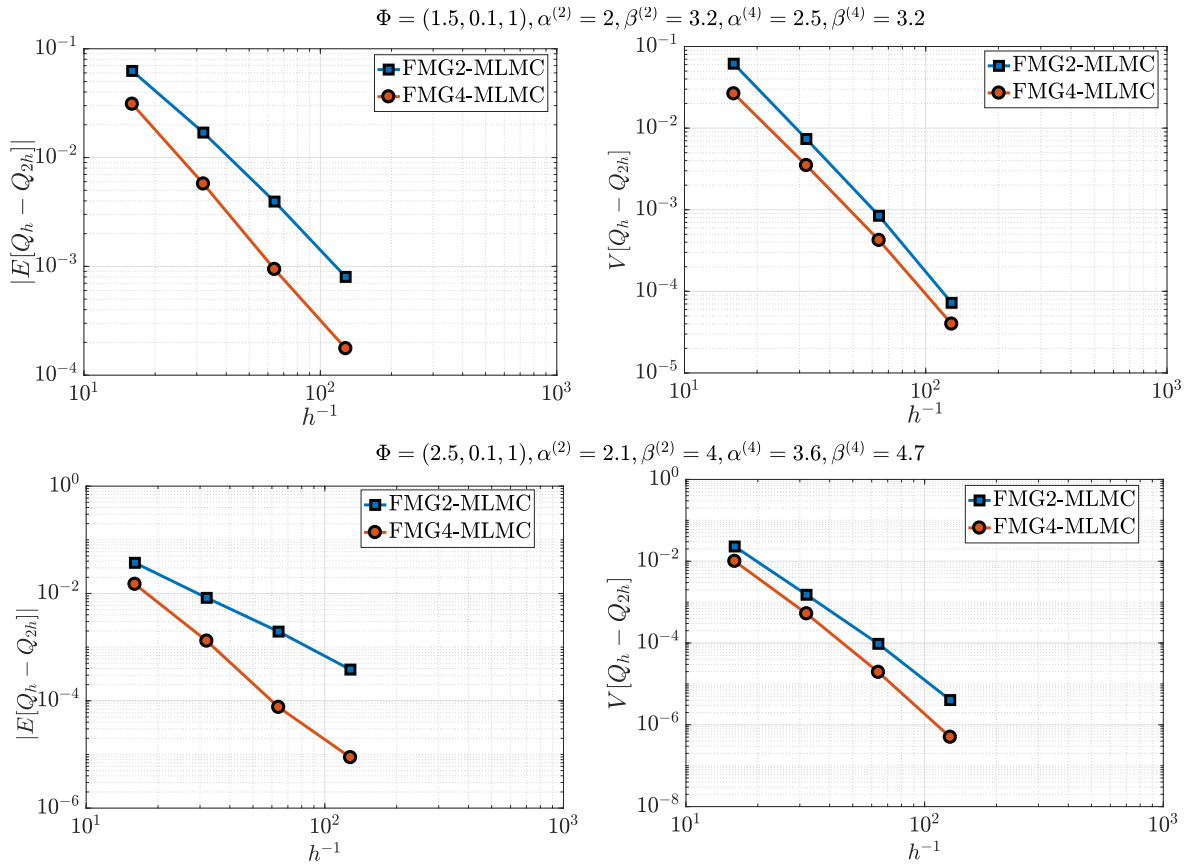
Figure 3.10: Comparison of parameter α *(left)* and β *(right)* observed for horizontal flux $-k\left.\frac{\partial p}{\partial x}\right|_{(\mathbf{x}^*)}$ where $\mathbf{x}^* = (1/2, 1/2)^T$ from FMG2-MLMC/FMG4-MLMC scheme.

with $h_0 = 1/4$, whereas the middle and the bottom two plots present more challenging parameters $\Phi = (1.5, 0.3, 3)$ and $\Phi = (2.5, 0.3, 3)$, respectively, both with $h_0 = 1/8$. For all these test cases, we observe similar improvements in the MLMC parameters as the previous quantity of interest from using the fourth-order discretization. Again, these parameters can be used to predict the asymptotic cost of different estimators. Here we would like to point out that, the discretization error from the second-order discretization is already very small ($< 10^{-3}$) on the coarser grids. For applications which do not require a very small tolerance, an estimator based on a fourth-order discretization may be more expensive.

In the presence of strong gradients, the constant term in the discretization error can be large. This is also observed in the top left plot of Figure 3.12. In these situations, using high-order schemes can also result in an expensive MLMC estimator. The issue can be tackled using a hybrid MLMC estimator which utilizes the second-order scheme on coarser levels and the fourth-order scheme on the finer levels. The cut-off level being the coarsest grid where the fourth-order solution becomes more accurate compared to the second-order solution. This approach is easy to implement and does not lead to any violation of the telescopic sum.

3.9. Conclusions

We have presented a multilevel estimator based on a fourth-order accurate discretization of the stochastic Darcy flow problem with smooth coefficient fields. Our goal was to exploit the additional regularity in the numerical solution of the PDE to achieve faster MLMC convergence. We utilized a fourth-order FV discretization scheme to approximate the PDE solution. Additionally, we provided
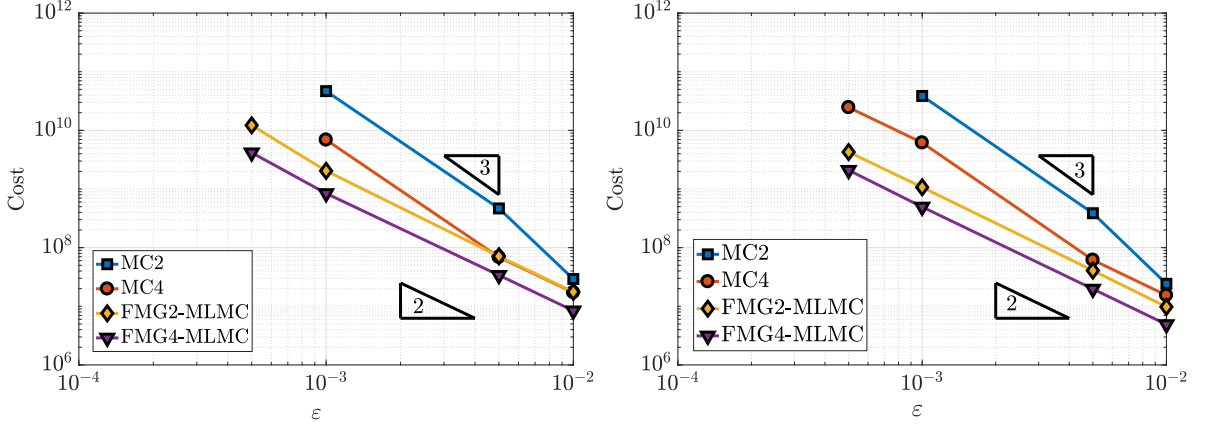
Figure 3.11: Cost to obtain RMSE less than ε in the horizontal flux for the 2D problem using different estimators. The covariance parameters are $\Phi = (1.5, 0.1, 1)$ (*left*) and $\Phi = (2.5, 0.1, 1)$ (*right*).

an efficient multigrid algorithm to solve the linear system arising from this discretization. A useful feature of this multigrid algorithm is that it uses a simple second-order multigrid solver combined with a defect correction strategy to obtain a fourth-order accurate solution. Further, we showed that this solver is able to handle coefficient fields with large variability and small correlation length scales. Numerical experiments clearly show the benefits of using a high-order discretization in terms of improved MLMC parameters α and β, which dictate the number of MLMC levels and samples, respectively, for smoothness $\nu_c > 1$. The fourth-order MLMC estimator reached the required tolerance at a much lower cost compared to the estimator based on the second-order discretization. Also, we showed that the fourth-order method leads to an asymptotic gain in case of the single-level MC estimator.

In this work, we confined ourselves to 2D problems and observed that $\beta > \gamma$ for all quantities of interest leading to an asymptotic cost $\mathcal{O}(\varepsilon^{-2})$. However, for 3D and unsteady problems for which γ is typically large, using a second-order method might lead to situations where $\beta < \gamma$ with the associated MLMC complexity of $\mathcal{O}(\varepsilon^{-2-(\gamma-\beta)/\alpha})$. In such scenarios, a higher α and β from the fourth-order scheme will improve this complexity.

Clearly, one of the issues with high-order schemes is that the cost of assembling and solving the linear system is more expensive compared to lower-order schemes (for same grid size) and therefore, the computational gains become evident only when the applications require a relatively small tolerance.

We have also proposed an approach to integrate a full multigrid solver with a multilevel estimator with the same mesh hierarchy. The algorithm is described in detail and all modifications required in the FMG solver are discussed. This combination results in computational savings which however, depend on the rate of decay of the MLMC samples with the levels. Although, we confined ourselves to a simplified version of the multigrid solver based on fixed transfer operators, this framework is easily extendable to more sophisticated solvers. Furthermore, the extension of this approach is straightforward to other uncertainty quantification problems in physics and engineering where a multigrid solver is used to solve the sparse linear system.
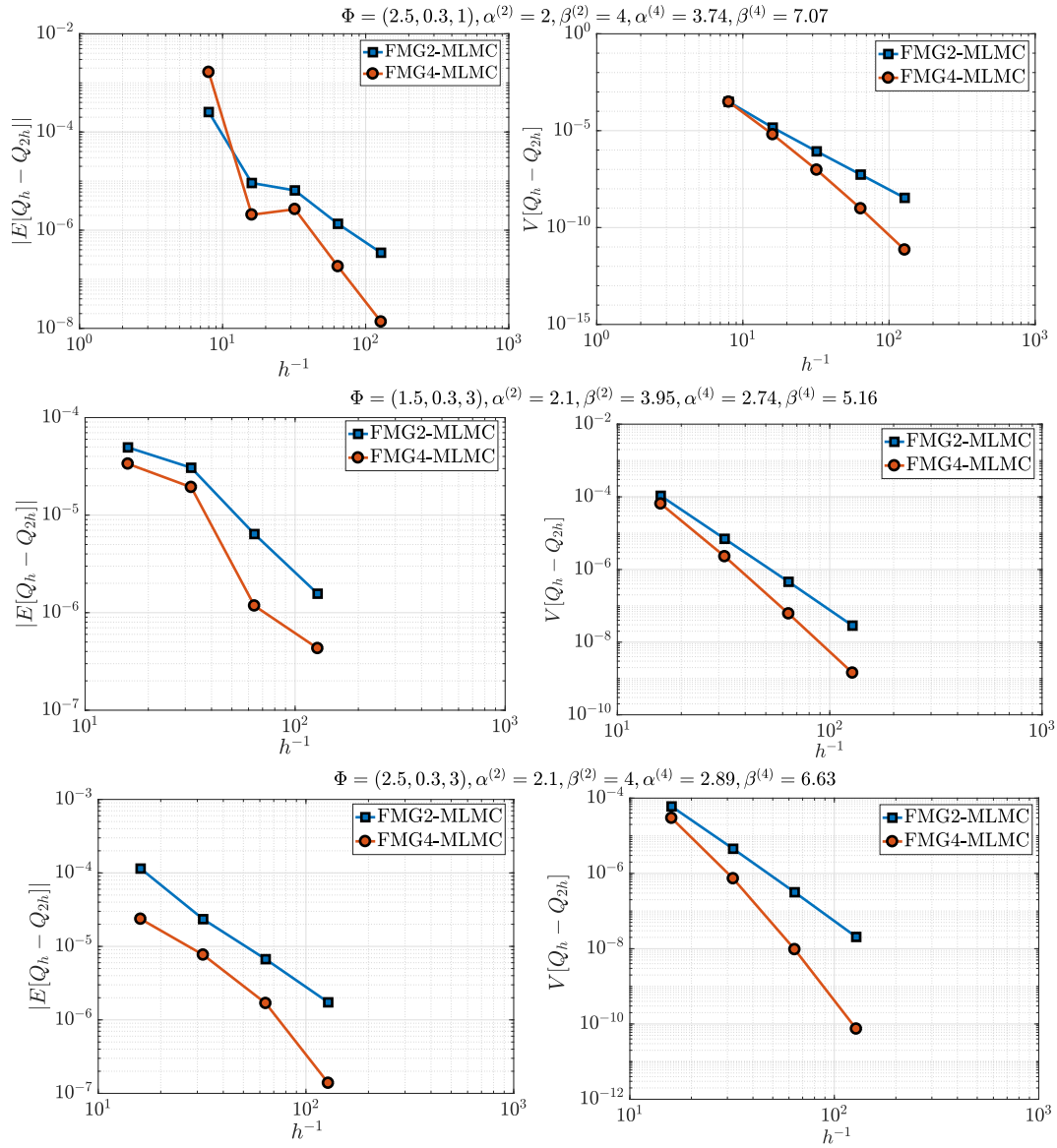
Figure 3.12: Comparison of parameter α *(left)* and β *(right)* observed for the mean of the solution $\frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} p d\mathbf{x}$ from FMG2-MLMC/FMG4-MLMC scheme.

Appendix

3.A. Approximation of high-order terms

Continuing the analysis from Section 3.3.2, we provide the approximation of the remaining high-order terms.

Computation of $\left\langle F^{BC} \right\rangle_{i+1/2,j}^{(4)}$

We now derive relations between face-averaged values and face-center point values. We can rewrite (3.15) as

$$\left\langle F^{BC} \right\rangle^{(4)} = k \left. \frac{\partial p}{\partial x} \right|_{i+\frac{1}{2},j} + \frac{h^2}{24} \left[\frac{\partial^2 k}{\partial y^2} \frac{\partial p}{\partial x} + 2 \frac{\partial k}{\partial y} \frac{\partial^2 p}{\partial y \partial x} + k \frac{\partial^3 p}{\partial y^2 \partial x} \right]_{i+\frac{1}{2},j} + \mathcal{O}(h^4). \qquad (3.60)$$

Thus, the face-averaged flux is expressed in terms of a point-wise value of k and $\frac{\partial p}{\partial x}$ and their derivatives at the center of the face. As the expansion in (3.15), one can also write

$$\langle k \rangle_{i+\frac{1}{2},j} = k_{i+\frac{1}{2},j} + \frac{h^2}{24} \frac{\partial^2 k}{\partial y^2}\bigg|_{i+\frac{1}{2},j} + \mathcal{O}(h^4), \tag{3.61}$$

$$\left\langle \frac{\partial p}{\partial x} \right\rangle_{i+\frac{1}{2},j} = \frac{\partial p}{\partial x}\bigg|_{i+\frac{1}{2},j} + \frac{h^2}{24} \frac{\partial^3 p}{\partial y^2 \partial x}\bigg|_{i+\frac{1}{2},j} + \mathcal{O}(h^4). \tag{3.62}$$

Hence, if we combine (3.61)-(3.62), we can reduce (3.60) to

$$\langle F^{BC} \rangle^{(4)}_{i+\frac{1}{2},j} = \langle k \rangle_{i+\frac{1}{2},j} \left\langle \frac{\partial p}{\partial x} \right\rangle_{i+\frac{1}{2},j} + \frac{h^2}{12} \frac{\partial k}{\partial y} \frac{\partial^2 p}{\partial y \partial x}\bigg|_{i+\frac{1}{2},j} + \mathcal{O}(h^4). \tag{3.63}$$

Computation of $\langle k \rangle_{i+1/2,j}$ **and** $\frac{\partial k}{\partial y}\big|_{i+1/2,j}$

Assuming that the coefficient field is smooth and is sampled at cell-centered locations, we first interpolate the coefficient values at face-centered locations using a centered fourth-order finite difference scheme, i.e.

$$k_{i+\frac{1}{2},j} = \frac{1}{16} \left[9(k_{i,j} + k_{i+1,j}) - (k_{i-1,j} + k_{i+2,j}) \right] + \mathcal{O}(h^4), \tag{3.64}$$

$$\frac{\partial^2 k}{\partial y^2}\bigg|_{i+\frac{1}{2},j} = \frac{1}{h^2} \left[k_{i+\frac{1}{2},j-1} - 2k_{i+\frac{1}{2},j} + k_{i+\frac{1}{2},j+1} \right] + \mathcal{O}(h^2). \tag{3.65}$$

We can compute $\langle k \rangle_{i+1/2,j}$ using (3.61). The term $\frac{\partial k}{\partial y}\big|_{i+1/2,j}$ from (3.63) only needs to be computed with $\mathcal{O}(h^2)$ accuracy as it is multiplied with h^2. Therefore,

$$\frac{\partial k}{\partial y}\bigg|_{i+\frac{1}{2},j} = \frac{1}{2h} [k_{i+\frac{1}{2},j+1} - k_{i+\frac{1}{2},j-1}] + \mathcal{O}(h^2). \tag{3.66}$$

3.B. Fourth-order linear operator

Next we describe the elements of the the fourth-order linear operator $\mathcal{L}_h^{(4)}$. For brevity, we use e, w, n and s to denote the four faces with centers $(x_{i+1/2}, y_j)$, $(x_{i-1/2}, y_j)$, $(x_i, y_{j+1/2})$ and $(x_i, y_{j-1/2})$, respectively. The fourth-order discrete operator away from the boundary can be represented as

$$\mathcal{L}_h^{(4)} = \frac{1}{1152h^2} \begin{bmatrix} 0 & l_{2,-1} & l_{2,0} & l_{2,1} & 0 \\ l_{1,-2} & l_{1,-1} & l_{1,0} & l_{1,1} & l_{1,2} \\ l_{0,-2} & l_{0,-1} & l_{0,0} & l_{0,1} & l_{0,2} \\ l_{-1,-2} & l_{-1,-1} & l_{-1,0} & l_{-1,1} & l_{-1,2} \\ 0 & l_{-2,-1} & l_{-2,0} & l_{-2,1} & 0 \end{bmatrix},$$

with

$$l_{0,0} = 1188 \sum_{d=e,w,n,s} \langle k \rangle_d,$$

$$l_{0,-1} = -44 \langle k \rangle_e - 1188 \langle k \rangle_w + 27 \sum_{d=n,s} (2 \langle k \rangle_d - \nabla^d k_d),$$

$$l_{0,1} = -44 \langle k \rangle_w - 1188 \langle k \rangle_e + 27 \sum_{d=n,s} (2 \langle k \rangle_d + \nabla^d k_d),$$

$$l_{-1,0} = -44 \langle k \rangle_n - 1188 \langle k \rangle_s + 27 \sum_{d=e,w} (2 \langle k \rangle_d - \nabla^d k_d),$$

$$l_{1,0} = -44 \langle k \rangle_s - 1188 \langle k \rangle_n + 27 \sum_{d=e,w} (2 \langle k \rangle_d + \nabla^d k_d),$$

$$l_{-1,-1} = \sum_{d=e,n} (2 \langle k \rangle_d - \nabla^d k_d) - 27 \sum_{d=w,s} (2 \langle k \rangle_d - \nabla^d k_d),$$

$$l_{1,1} = \sum_{d=w,s} (2 \langle k \rangle_d + \nabla^d k_d) - 27 \sum_{d=e,n} (2 \langle k \rangle_d + \nabla^d k_d),$$

$$l_{1,-1} = -27(2 \langle k \rangle_w + \nabla^w k_w) - (2 \langle k \rangle_s - \nabla^s k_s) - (2 \langle k \rangle_e + \nabla^e k_e) - 27(2 \langle k \rangle_n - \nabla^n k_n),$$

$$l_{-1,1} = -(2 \langle k \rangle_w - \nabla^w k_w) - 27(2 \langle k \rangle_s + \nabla^s k_s) - 27(2 \langle k \rangle_e - \nabla^e k_e) - (2 \langle k \rangle_n + \nabla^n k_n),$$

$$l_{0,-2} = 44 \langle k \rangle_w, \qquad l_{0,2} = 44 \langle k \rangle_e, \qquad l_{-2,0} = 44 \langle k \rangle_s, \qquad l_{2,0} = 44 \langle k \rangle_n,$$

$$l_{-2,-1} = 2 \langle k \rangle_s - \nabla^s k_s, \qquad l_{-2,1} = 2 \langle k \rangle_s + \nabla^s k_s,$$

$$l_{2,-1} = 2 \langle k \rangle_n - \nabla^n k_n, \qquad l_{2,1} = 2 \langle k \rangle_n + \nabla^n k_n,$$

$$l_{-1,-2} = 2 \langle k \rangle_w - \nabla^w k_w, \qquad l_{-1,2} = 2 \langle k \rangle_e - \nabla^e k_e,$$

$$l_{1,-2} = 2 \langle k \rangle_w + \nabla^w k_w, \qquad l_{1,2} = 2 \langle k \rangle_e + \nabla^e k_e.$$

where $\nabla^d k_d$ denote the gradient of k along the face d and is computed using Eq. (3.66).

4

Local Fourier analysis of multigrid methods for SPDEs

A non-standard Local Fourier Analysis (LFA) variant for accurately predicting the multigrid convergence of problems with random and jumping coefficients is discussed. This LFA method is based on a specific basis of the Fourier space rather than the commonly used Fourier modes. To show the utility of this analysis, as an example, a simple cell-centered multigrid method for solving a steady-state single phase flow problem in a random porous medium is analyzed. A number of challenging benchmark problems are considered to demonstrate the prediction capability of the proposed LFA. The information provided by this analysis can help us to estimate apriori the time needed for solving certain uncertainty quantification problems by means of a multigrid multilevel Monte Carlo method.

4.1. Introduction

For any sampling-based UQ technique such as Monte Carlo type methods, the availability of a highly efficient and robust (w.r.t. the random inputs) iterative solver becomes critical. In general, the sample-wise computational cost can become highly heterogeneous, depending on the random inputs. Therefore, if the performance statistics of such solvers were known apriori, one could utilize this information to optimize and parallelize the MC simulations efficiently.

In this chapter, we present a non-standard Local Fourier Analysis (LFA) technique to predict the convergence rate of multigrid solvers for problems involving random and jumping coefficients. Standard LFA techniques are typically based on constant coefficient discretization stencils, whereas for stochastic PDEs we encounter varying coefficients throughout the computational domain, due to the randomness. One of the main contributions of this work is to generalize the LFA towards problems with random and jumping coefficients, with the aim of predicting, apriori, the total time needed to solve UQ problems. Some efforts have already been done in [67] regarding the generalization of LFA for jumping coefficients. The novelty of our approach lies in the choice of basis functions. Here, we utilize a new basis from the Fourier space rather than the standard Fourier modes. We benchmark the prediction capability of the proposed LFA technique using a set of challenging jumping coefficient problems and a number of spatially correlated random fields with varying heterogeneity.

Another highlight of this chapter is the use of the LFA predictions to optimize the multilevel Monte Carlo computations. Combining MLMC methods with multigrid seems natural, where the multigrid method is employed for the numerical solution of a PDE which is based on a sample of the stochastic quantity on a certain (fine or coarse) scale (see also the previous chapter). Due to the stochasticity, however, we deal with PDE problems with jumping coefficients, where different jump patterns are encountered each time a new random field is generated. The generalization of the LFA towards these PDEs will provide us insight in the average number of multigrid iterations, the spread of the convergence factors, amongst other things. This information helps to estimate the total CPU time needed for the multiple multigrid computations in a multilevel Monte Carlo setting. We consider it useful to construct a technique to assess the quality of the choice of the multigrid components in the context of the PDEs with random problems before the actual multigrid computation has taken place.

In this work, we will employ a basic cell-centered multigrid (CCMG) algorithm for solving elliptic PDEs with a variable coefficient field. The components of this algorithm include a simple Gauss-Seidel iteration as the smoother, a piecewise constant prolongation operator and its adjoint as the restriction and a direct discretization technique to define the discrete operators on the coarse grids. We show that for this special combination, the coarse grid discretization operators are equivalent to the ones obtained from commonly used Galerkin operators [16]. We utilize this CCMG method to perform MLMC simulations with different permeability parameters. It may be surprising that such a basic algorithm converges well in the context of the generated random fields, where computation takes place for thousands of different samples. Despite we restrict ourselves to this basic CCMG to demonstrate the accuracy of the predictions of the novel LFA technique, we emphasize that this approach can be used for a wider range of problems, discretizations, and multigrid methods. In fact, we think that the proposed LFA allows us to deal in an easy way with challenging problems for which a standard LFA is very difficult to apply or even impracticable.

The chapter is organized as follows. In Section 4.2 we introduce the context of PDEs with jumping and random coefficients, together with their discretization by a cell-centered finite volume scheme. A discussion on multigrid methods for this type of problems is also included, and the multigrid components that will be considered in this work are also defined. Section 4.3 is devoted to the generalization of the LFA to deal with jumping coefficients and problems with random fields. In Section 4.4, we present results obtained by this analysis for different benchmark problems with jumping coefficients. Section 4.5 presents the LFA results for problems with random coefficients, and in Section 4.7 multilevel Monte Carlo computations for PDEs with random coefficients are presented. Finally, in Section 4.8 conclusions are drawn.

4.2. Jumping coefficients, random coefficients, multigrid

Robust and efficient iterative solution methods are very relevant for partial differential equations with variable coefficients. For PDEs with jumping coefficients, multigrid methods have already shown to be this type of solvers. When using Monte Carlo methods, in the case of elliptic PDEs with random coefficient fields, many samples of the random field are generated and for each field, the numerical solution should be computed. This can take substantial CPU time if very many samples are required. For a fixed sample of the random field, we deal with an elliptic PDE with varying coefficients, due to the randomness. Multigrid comes in naturally as a highly efficient solution method for the resulting PDEs. In this section, we explain this setting and we briefly describe an efficient multigrid method based on a cell-centered grid and a finite volume discretization.

4.2.1 PDEs with jumping and with random coefficients

We start with classical PDE problems with jumping coefficients. In particular, we deal with the following two-dimensional diffusion equation on the square domain $\mathcal{D} = (0, \ell)^2$,

$$-\nabla \cdot (k(\mathbf{x})\nabla p(\mathbf{x})) = f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{D}, \tag{4.1}$$
$$p(\mathbf{x}) = g(\mathbf{x}), \quad \mathbf{x} \in \partial\mathcal{D}, \tag{4.2}$$

where $k(\mathbf{x})$ is a function which may be discontinuous across internal boundaries.

To discretize this problem, we use a *cell-centered finite volume method* based on the harmonic average of the diffusion coefficient $k(\mathbf{x})$. We consider a uniform grid $\mathcal{D}_h$ with the same step size $h = \ell/M, M \in \mathbb{N}$ in both directions,

$$\mathcal{D}_h = \{(x_{i_1}, x_{i_2}); x_{i_\alpha} = (i_\alpha - 1/2)h, i_\alpha = 1, \dots, M, \alpha = 1, 2\}. \tag{4.3}$$

This gives, for each interior cell with center (x_{i_1}, x_{i_2}), denoted by $\mathcal{D}_h^{i_1, i_2}$, a five-point scheme

$$c_{i_1,i_2}^h p_{i_1,i_2} + w_{i_1,i_2}^h p_{i_1-1,i_2} + e_{i_1,i_2}^h p_{i_1+1,i_2} + s_{i_1,i_2}^h p_{i_1,i_2-1} + n_{i_1,i_2}^h p_{i_1,i_2+1} = f_{i_1,i_2}^h, \tag{4.4}$$

where

$$
\begin{aligned}
w_{i_1,i_2}^h &= -\frac{2}{h^2} \frac{k_{i_1,i_2} k_{i_1-1,i_2}}{k_{i_1,i_2} + k_{i_1-1,i_2}}, \quad e_{i_1,i_2}^h = -\frac{2}{h^2} \frac{k_{i_1,i_2} k_{i_1+1,i_2}}{k_{i_1,i_2} + k_{i_1+1,i_2}}, \\
s_{i_1,i_2}^h &= -\frac{2}{h^2} \frac{k_{i_1,i_2} k_{i_1,i_2-1}}{k_{i_1,i_2} + k_{i_1,i_2-1}}, \quad n_{i_1,i_2}^h = -\frac{2}{h^2} \frac{k_{i_1,i_2} k_{i_1,i_2+1}}{k_{i_1,i_2} + k_{i_1,i_2+1}}, \\
c_{i_1,i_2}^h &= -(w_{i_1,i_2}^h + e_{i_1,i_2}^h + n_{i_1,i_2}^h + s_{i_1,i_2}^h),
\end{aligned}
$$

with, for instance, k_{i_1,i_2} the diffusion coefficient associated with the cell $\mathcal{D}_h^{i_1,i_2}$. By *interior cell* we mean a cell for which none of its edges lies at the boundary of the domain. This scheme is changed appropriately for the cells close to the boundary.

We also recall the elliptic PDEs with random coefficient fields from Chapter 3. The PDE of our interest describes the steady-state single-phase flow in a random porous medium. Denoting by ω an event in the probability space $(\Omega, \mathbb{F}, \mathbb{P})$, with sample space Ω, σ-algebra $\mathbb{F}$ and probability measure $\mathbb{P}$, the permeability in the porous medium is described by $k(\mathbf{x}, \omega) : \overline{\mathcal{D}} \times \Omega \to \mathbb{R}_+$. The PDE is then given by

$$-\nabla \cdot (k(\mathbf{x}, \omega)\nabla p(\mathbf{x}, \omega)) = f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{D}, \omega \in \Omega, \tag{4.5}$$

with f as a source term. The engineering interest in the solution of (4.5) is typically found in expected values of linear functionals of the solution p, denoted by $Q := Q(p)$.

To discretize these problems, we use the same cell-centered finite volume method based on the harmonic average of the random diffusion coefficient as previously described for problems with jumping coefficients. We make the common assumption that the permeability random field is constant over each cell of the grid.

4.2.2 Multigrid for PDEs with jumping and with random coefficients

In this work, the multigrid components for the above cell-centered discrete problems are chosen as follows. We use a lexicographic Gauss-Seidel iteration as the multigrid smoother and we consider standard coarsening obtained by doubling the mesh size in both directions. Each coarse cell is the union of four fine cells, and, since the unknowns are located in the cell-centers, this results

in a non-nested hierarchy of grids. We consider a simple prolongation operator P_{2h}^h, that is, the piecewise constant interpolation operator. In stencil notation, it is given by

$$
P_{2h}^h = \left.\begin{bmatrix} 1 & & 1 \\ & \star & \\ 1 & & 1 \end{bmatrix}\right._{2h}^{h}, \tag{4.6}
$$

where $\star$ denotes the position of a coarse grid unknown. The classical stencil notation shows the contribution of the coarse grid node to the neighbouring fine grid nodes. The restriction operator R_h^{2h} is chosen as the scaled adjoint of the prolongation, given in stencil form by

$$
R_h^{2h} = \frac{1}{4}\left.\begin{bmatrix} 1 & & 1 \\ & \star & \\ 1 & & 1 \end{bmatrix}\right._{h}^{2h}. \tag{4.7}
$$

The coarse grid operators are constructed by direct discretization defining the diffusion coefficients at the edges of the coarse cells appropriately, which we will describe in more detail. We assume that the diffusion coefficient $k(\mathbf{x})$ is piecewise constant on the fine grid. The flux over an edge, dependent on the solution in the two adjacent cells, is calculated based on the harmonic average. The values of the diffusion coefficients at a coarse edge located between two coarse cells, however, are calculated as the arithmetic average of the corresponding fine grid coefficients, see Figure 4.1 for a more detailed description. As it was pointed in [68], this direct discretization procedure is equivalent to the
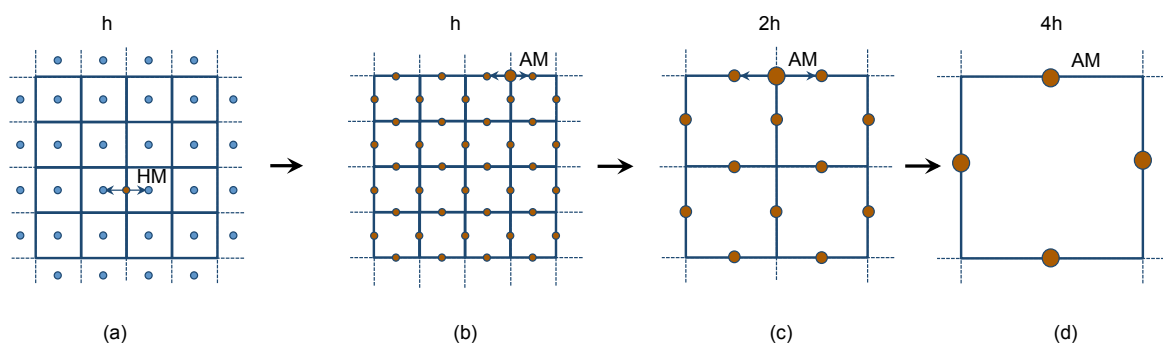

Figure 4.1: Schematic representation of permeability upscaling used in the multigrid hierarchy (h-2h-4h). (a) Permeability values generated at cell-centers (blue dots). (b) Permeability values at face-centers (red dots) obtained from the harmonic mean (HM) of permeabilities from two adjacent cell-centers. (c) Permeability at face-centers (bigger red dots) of 2h-grid is the arithmetic mean (AM) of permeabilities from face-centers of the h-grid. (d) Permeability at face-centers (biggest red dots) of 4h-grid is the arithmetic mean (AM) of permeabilities from face-centers of the 2h-grid.

often used Galerkin approach, i.e., $L_{2h} = \frac{1}{2}R_h^{2h}L_h P_{2h}^h$, but computationally more efficient. The factor $1/2$ in the previous expression is due to the lack of consistency of the operator $R_h^{2h}L_h P_{2h}^h$ with the differential operator [69]. In the next result, we prove that both discretizations are indeed equivalent.

PROPOSITION 1. *Let L_h be the fine-grid operator based on the cell-centered finite volume discretization of problem (4.5) on a uniform grid of mesh size $h = \ell/M$ with M even. Let P_{2h}^h be the piecewise constant prolongation operator and R_h^{2h} its adjoint. Then, the Galerkin coarse grid operator $L_{2h} = \frac{1}{2}R_h^{2h}L_h P_{2h}^h$ is equivalent to a direct discretization on the coarse grid based on the arithmetic average of the corresponding fine grid coefficients.*

Proof: We prove the equivalence for a coarse grid cell $\mathcal{D}_{2h}^{i_1,i_2}$ such that none of its edges lies on the boundary of the domain. The equivalence for coarse cells close to boundaries with Dirichlet or Neumann boundary conditions can be proven similarly. By applying the restriction operator R_h^{2h} in (4.7), the equation associated with the cell $\mathcal{D}_{2h}^{i_1,i_2}$ by using the Galerkin approach is given by

$$\frac{1}{2}(R_h^{2h}L_hP_{2h}^h p)_{i_1,i_2} = \frac{1}{8}((L_hP_{2h}^h p)_{2i_1,2i_2} + (L_hP_{2h}^h p)_{2i_1-1,2i_2} + (L_hP_{2h}^h p)_{2i_1,2i_2-1} + (L_hP_{2h}^h p)_{2i_1-1,2i_2-1}). \tag{4.8}$$

Taking into account that the prolongation operator is piecewise constant, we obtain

$$(L_hP_{2h}^h p)_{2i_1,2i_2} = e_{2i_1,2i_2}^h p_{i_1+1,i_2} + w_{2i_1,2i_2}^h p_{i_1,i_2} + n_{2i_1,2i_2}^h p_{i_1,i_2+1} + s_{2i_1,2i_2}^h p_{i_1,i_2} + c_{2i_1,2i_2}^h p_{i_1,i_2},$$

and similar expressions for the other terms in (4.8). By substituting these expressions in (4.8), the following discretization for the coarse cell $\mathcal{D}_{2h}^{i_1,i_2}$ is obtained

$$\frac{1}{2}(R_h^{2h}L_hP_{2h}^h p)_{i_1,i_2} = c_{i_1,i_2}^{2h} p_{i_1,i_2} + w_{i_1,i_2}^{2h} p_{i_1-1,i_2} + e_{i_1,i_2}^{2h} p_{i_1+1,i_2} + s_{i_1,i_2}^{2h} p_{i_1,i_2-1} + n_{i_1,i_2}^{2h} p_{i_1,i_2+1}, \tag{4.9}$$

where

$$w_{i_1,i_2}^{2h} = \frac{1}{8}\left(w_{2i_1-1,2i_2}^h + w_{2i_1-1,2i_2-1}^h\right), \quad e_{i_1,i_2}^{2h} = \frac{1}{8}\left(e_{2i_1,2i_2}^h + e_{2i_1,2i_2-1}^h\right),$$

$$s_{i_1,i_2}^{2h} = \frac{1}{8}\left(s_{2i_1,2i_2-1}^h + s_{2i_1-1,2i_2-1}^h\right), \quad n_{i_1,i_2}^{2h} = \frac{1}{8}\left(n_{2i_1-1,2i_2}^h + n_{2i_1,2i_2}^h\right),$$

$$c_{i_1,i_2}^{2h} = -(w_{i_1,i_2}^{2h} + e_{i_1,i_2}^{2h} + n_{i_1,i_2}^{2h} + s_{i_1,i_2}^{2h}).$$

We observe that this scheme is equivalent to a direct discretization on the coarse grid where the diffusion coefficients on the edges are the arithmetic averages of the corresponding fine grid coefficients.

Remark 4.2.1 *In order to achieve a mesh-independent multigrid convergence following the analysis from [70], the next condition must be satisfied:*

$$m_p + m_r > M_{pde}, \tag{4.10}$$

where the orders m_p and m_r are the highest degree plus one of the polynomials that are exactly interpolated by P_{2h}^h and R_h^{2h}, respectively, and M_{pde} is the order of the PDE to be solved. For PDE (4.5), we have $M_{pde} = 2$ and for the considered operators (4.6) and (4.7), we get $m_p = m_r = 1$, which does not satisfy the inequality (4.10). In [71] it is shown, however, that this condition is not needed to prove uniform convergence. The idea is to use the power of the theoretical approach provided in [72].

4.2.3   Discussion about other multigrid methods for jumping coefficients

In the context of *algebraic multigrid* methods for the numerical solution of partial differential equations, basically, two prevailing methods have proved their use for multiple engineering problems, i.e., algebraic multigrid and aggregation-based multigrid methods, [73–77]. These methods converge remarkably well, for example, for scalar PDEs with jumping coefficients. It is not always easily understood why these methods, and particularly the aggregation-based method, converge so well.

The origin of these algebraic methods may be found already in the early days of multigrid, where black-box multigrid with operator-dependent transfer operators (restriction and prolongation) and

Galerkin coarse grid operators for structured vertex-centered Cartesian grids was proposed in [78–80]. This can be seen as a predecessor of classical AMG, where these components were essentially enhanced by a flexible coarsening strategy.

The aggregation-based multigrid methods, with their origin in the work by Mandel [77] (smoothed aggregation), may be related to the cell-centered multigrid methods as proposed in [54, 81]. In [81], it was shown that constant, i.e., operator-*independent*, transfer operators, in combination with Galerkin coarse grid discretization provided highly efficient multigrid results for cell-centered discretizations of elliptic PDEs that included jumping coefficients. These cell-centered multigrid components were augmented with robust smoothing, like Incomplete Lower-Upper decomposition (ILU) relaxation. The individual contributions of the coarse grid correction and the smoothing parts were difficult to distinguish. Also, a cell-centered multigrid based on coarsening by a factor of three together with operator-dependent interpolations was explored in [82].

4.3. Local Fourier Analysis for variable coefficients

In this section, we describe LFA in a setting which allows us to estimate the multigrid convergence factors for problems with jumping coefficients and problems with random fields. A discrete linear operator with constant coefficients, which is formally defined on an infinite grid, is usually assumed for carrying out a standard local Fourier analysis. As we will show, this assumption can be relaxed by considering a discrete operator with constant coefficients *in appropriate infinite subgrids*. This allows us to generalize the analysis to problems for which the discrete operator consists of different stencils. A key point in this improved analysis is to consider a specific basis of the Fourier space, rather than the standard basis which is based on the Fourier modes. The use of this new basis will simplify the analysis.

We start from a regular infinite grid $\mathcal{D}_h$ with grid size h in both directions. Such an infinite grid will be *split into $n \times n$ subgrids* in the following way. First of all, a window comprising $n \times n$ cells of the original grid is *adequately* chosen, and, subsequently, we consider its periodic extension. The choice of the size of the $n \times n$ window is made such that the variability of the discrete operator in the computational grid can be appropriately represented, as will be explained by means of examples of different nature. Once n is fixed, the infinite subgrids are defined as follows (see Figure 4.2 for an example with $n = 2$),

$$\mathcal{D}_h^{kl} = \{(k,l)h + (nk_1, nk_2)h \mid k_1, k_2 \in \mathbb{Z}\}, \ k, l = 0, \dots, n-1. \tag{4.11}$$

For each low frequency, $\theta^{00} \in \Theta_{nh} = (-\pi/nh, \pi/nh]^2$, we introduce the grid-functions:
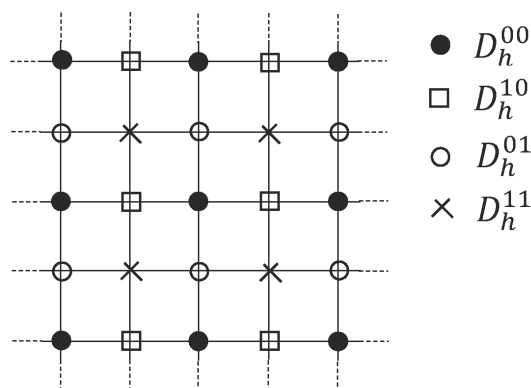

Figure 4.2: Infinite grid $\mathcal{D}_h$ divided into the corresponding subgrids for $n = 2$.

$$\psi_h^{kl}(\theta^{00}, \mathbf{x}) = \varphi_h(\theta^{00}, \mathbf{x})\chi_{\mathcal{D}_h^{kl}}(\mathbf{x}), \; k, l = 0, \dots, n-1, \; \mathbf{x} \in \mathcal{D}_h, \tag{4.12}$$

where $\varphi_h(\theta^{00}, \mathbf{x}) = e^{\iota\theta^{00} \cdot \mathbf{x}}$ is the standard Fourier mode on $\mathcal{D}_h$ corresponding to the frequency θ^{00}. It is easy to see that the subspace generated by these n^2 grid-functions,

$$\mathcal{F}_h^{n^2}(\theta^{00}) = span\{\psi_h^{kl}(\theta^{00}, \cdot), \; k, l = 0, \dots, n-1\} \tag{4.13}$$

is the same as the one spanned by the n^2 Fourier modes $\varphi_h(\theta_{kl}^{00}, \cdot)$ associated with the frequencies:

$$\theta_{kl}^{00} = \theta^{00} + (k, l)\frac{2\pi}{nh}, \; k, l = 0, \dots, n-1. \tag{4.14}$$

In the case $n = 2$, the basis $\{\psi_h^{00}(\theta^{00}, \cdot), \psi_h^{11}(\theta^{00}, \cdot), \psi_h^{10}(\theta^{00}, \cdot), \psi_h^{01}(\theta^{00}, \cdot)\}$ is related to the standard basis of Fourier modes $\{\varphi_h(\theta^{00}, \cdot), \varphi_h(\theta^{11}, \cdot), \varphi_h(\theta^{10}, \cdot), \varphi_h(\theta^{01}, \cdot)\}$ in the following way:

$$\begin{pmatrix} \psi_h^{00}(\theta^{00}, \cdot) \\ \psi_h^{11}(\theta^{00}, \cdot) \\ \psi_h^{10}(\theta^{00}, \cdot) \\ \psi_h^{01}(\theta^{00}, \cdot) \end{pmatrix} = \mathcal{M} \begin{pmatrix} \varphi_h(\theta^{00}, \cdot) \\ \varphi_h(\theta^{11}, \cdot) \\ \varphi_h(\theta^{10}, \cdot) \\ \varphi_h(\theta^{01}, \cdot) \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} \varphi_h(\theta^{00}, \cdot) \\ \varphi_h(\theta^{11}, \cdot) \\ \varphi_h(\theta^{10}, \cdot) \\ \varphi_h(\theta^{01}, \cdot) \end{pmatrix}. \tag{4.15}$$

It is well-known that Fourier modes are eigenfunctions of any constant coefficient linear discrete operator L_h, that is, $L_h\varphi_h(\theta, \mathbf{x}) = \widetilde{L}_h(\theta)\varphi_h(\theta, \mathbf{x})$. Therefore, the representation of L_h with respect to a basis of n^2 Fourier modes is a diagonal matrix with diagonal elements $\widetilde{L}_h(\theta^{kl})$ with $k, l = 0, \dots, n-1$. In general, the Fourier representation with respect to the basis of functions $\{\psi_h^{kl}\}_{k,l=0}^{n-1}$ is a dense matrix. We will denote it by $\widehat{L}_h(\theta^{00})$.

If we consider the five-point standard discretization of the Laplace operator on a uniform grid of mesh size h,

$$\frac{1}{h^2} \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}, \tag{4.16}$$

its Fourier symbol with respect to the standard basis of Fourier modes is a diagonal matrix with diagonal elements equal to

$$\frac{1}{h^2}(4 - 2\cos(\theta_x^{kl}) - 2\cos(\theta_y^{kl})),$$

(see [14], for instance) whereas the Fourier representation with respect to the new basis in the case $n = 2$ is given by

$$\widehat{L}_h(\theta^{00}) = \frac{2}{h^2} \begin{pmatrix} 2 & 0 & \cos(\theta_x^{00}) & \cos(\theta_y^{00}) \\ 0 & 2 & \cos(\theta_y^{00}) & \cos(\theta_x^{00}) \\ \cos(\theta_x^{00}) & \cos(\theta_y^{00}) & 2 & 0 \\ \cos(\theta_y^{00}) & \cos(\theta_x^{00}) & 0 & 2 \end{pmatrix}, \text{ with } \theta^{00} = (\theta_x^{00}, \theta_y^{00}). \tag{4.17}$$

Notice that, for example, the first row of the previous symbol is obtained by looking at the decomposition of the stencil (4.16) into the connections among the unknowns located at the different subgrids $\mathcal{D}_h^{kl}$ defined in (4.11). In particular, following the notations in Figure 4.2, the •-•, • − □ and • − ∘ connections are given by the following stencils

$$\frac{1}{h^2}[4], \quad \frac{1}{h^2}\begin{bmatrix} -1 & \bullet & -1 \end{bmatrix}, \quad \frac{1}{h^2}\begin{bmatrix} -1 \\ \bullet \\ -1 \end{bmatrix},$$

giving rise to the symbols

$$\frac{4}{h^2}, \quad \frac{2}{h^2}\cos(\theta_x^{00}), \quad \frac{2}{h^2}\cos(\theta_y^{00}),$$

which appear in the first row of (4.17), whereas there is no $\bullet - \times$ connection. The rest of the rows is analogously computed. The procedure to obtain the Fourier symbol of a smoothing operator S_h, which is based on a splitting of the discrete operator $L_h = L_h^+ + L_h^-$, is analogous with the new basis. The smoothing iteration is given by

$$L_h^+ \overline{w}_h + L_h^- w_h = f_h,$$

with w_h the approximation of the solution before the smoothing step and $\overline{w}_h$ the approximation after the smoothing step. By computing the symbols of L_h^+ and L_h^- as before, the Fourier symbol of the smoothing operator is given by

$$\widehat{S}_h(\theta^{00}) = -(\widehat{L}_h^+)^{-1}(\theta^{00})\widehat{L}_h^-(\theta^{00}). \tag{4.18}$$

The Fourier symbol corresponding to a lexicographic Gauss-Seidel iteration for the five-point standard discretization of the Laplace operator on a uniform grid of mesh size h, in the case $n = 2$, is as in (4.18), where

$$\widehat{L}_h^+(\theta^{00}) = \frac{2}{h^2}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & \cos(\theta_y^{00}) & \cos(\theta_x^{00}) \\ \cos(\theta_x^{00}) & 0 & 2 & 0 \\ \cos(\theta_y^{00}) & 0 & 0 & 2 \end{pmatrix}, \quad \widehat{L}_h^-(\theta^{00}) = \widehat{L}_h(\theta^{00}) - \widehat{L}_h^+(\theta^{00}). \tag{4.19}$$

Once the Fourier representation of the smoothing operator with respect to the new basis is obtained, we can define the smoothing factor by using the change of basis matrix. For example, for the case $n = 2$, the smoothing factor is obtained by

$$\mu(S_h) = \sup_{\theta^{00} \in \Theta_{2h}} \rho(\mathcal{Q}_h \mathcal{M} \widehat{S}_h(\theta^{00})\mathcal{M}^{-1}),$$

where $\mathcal{Q}_h$ is the projection operator onto the space of high-frequency components and $\mathcal{M}$ is the change of basis matrix given in (4.15).

4.3.1   LFA formulations for cell-centered grids

In Section 4.2, we didn't distinguish between cell- and vertex-centered grids. The generalized LFA indeed works well for both types of discretization. By introducing the coarse grids and their relation with the fine grids, we need to fix the approach of interest. Since here we will focus on cell-centered discretizations, from now on the description of the analysis will be given for this case, although it may be applied to the vertex-centered case in a similar way by defining appropriately the coarse meshes.

According to the location of the coarse-grid points in a regular cell-centered grid, we define for a fixed n the following infinite coarse subgrids of $\mathcal{D}_{2h}$:

$$\mathcal{D}_{2h}^{kl} = \{(h/2, h/2) + (k,l)2h + (nk_1, nk_2)h \mid k_1, k_2 \in \mathbb{Z}\}, \ k,l = 0, \dots, n/2 - 1. \tag{4.20}$$

Due to the relation between the grid-functions of the new Fourier basis given in (4.12) and the standard Fourier modes, it can be shown that the coarse-grid correction operator $C_h = I_h -$

$P_{2h}^h L_{2h}^{-1} R_h^{2h} L_h$, where P_{2h}^h and R_h^{2h} are the prolongation and restriction operators, L_h and L_{2h} are the fine- and coarse-grid operators and I_h is the identity, satisfies the following invariance property:

$$C_h : \mathcal{F}_h^{n^2}(\theta^{00}) \to \mathcal{F}_h^{n^2}(\theta^{00}).$$

More concretely, for $\theta^{00} \in \Theta_{2nh} = (-\pi/2nh, \pi/2nh]^2$, the following properties of the operators in C_h are fulfilled:

1. $L_h, I_h : \mathcal{F}_h^{n^2}(\theta^{00}) \to \mathcal{F}_h^{n^2}(\theta^{00})$,

2. $L_{2h} : \mathcal{F}_{2h}^{n^2/4}(2\theta^{00}) \to \mathcal{F}_{2h}^{n^2/4}(2\theta^{00})$,

3. $R_h^{2h} : \mathcal{F}_h^{n^2}(\theta^{00}) \to \mathcal{F}_{2h}^{n^2/4}(2\theta^{00})$,

4. $P_{2h}^h : \mathcal{F}_{2h}^{n^2/4}(2\theta^{00}) \to \mathcal{F}_h^{n^2}(\theta^{00})$.

From these invariance properties we can compute the Fourier representations of the corresponding operators. As an example, we will describe next the representation of R_h^{2h} with respect to the grid-functions $\{\psi_{2h}^{kl}(2\theta^{00})\}_{k,l=0}^{n/2-1}$ and $\{\psi_h^{kl}(\theta^{00})\}_{k,l=0}^{n-1}$, for the restriction operators considered in this work.

We first consider the basic restriction operator obtained as the adjoint of the piecewise constant prolongation operator with stencil form (4.7). Its Fourier representation with respect to the new Fourier basis is given by

$$\widehat{R}_h^{2h}(\theta^{00}) = \frac{1}{4} \left(\; e^{-\iota(\theta_x^{00}+\theta_y^{00})/2} \quad e^{\iota(\theta_x^{00}+\theta_y^{00})/2} \quad e^{\iota(\theta_x^{00}-\theta_y^{00})/2} \quad e^{\iota(\theta_y^{00}-\theta_x^{00})/2} \; \right).$$

In the case of the cell-centered restriction operator by Wesseling/Khalil [54], that is,

$$R_h^{2h} = \frac{1}{16} \begin{bmatrix} 1 & 1 & & 0 & 0 \\ 1 & 3 & & 2 & 0 \\ & & \star & & \\ 0 & 2 & & 3 & 1 \\ 0 & 0 & & 1 & 1 \end{bmatrix}_h^{2h}, \tag{4.21}$$

the Fourier representation is given by

$$\widehat{R}_h^{2h}(\theta^{00}) = \frac{1}{16} \begin{pmatrix} e^{-\iota(\theta_x^{00}+\theta_y^{00})/2}(2 + e^{2\iota\theta_y^{00}} + e^{2\iota\theta_x^{00}}) \\ e^{\iota(\theta_x^{00}+\theta_y^{00})/2}(2 + e^{-2\iota\theta_y^{00}} + e^{-2\iota\theta_x^{00}}) \\ e^{\iota(\theta_x^{00}-\theta_y^{00})/2}(3 + e^{2\iota(\theta_y^{00}-\theta_x^{00})}) \\ e^{\iota(\theta_y^{00}-\theta_x^{00})/2}(3 + e^{2\iota(\theta_x^{00}-\theta_y^{00})}). \end{pmatrix}^T.$$

As an immediate consequence of these invariance properties and the invariance property of the smoothing operator, also the two-grid operator $K_h^{2h} = C_h S_h^\nu$, where ν denotes the number of smoothing steps, leaves the subspaces $\mathcal{F}_h^{n^2}(\theta^{00})$ invariant. Its Fourier representation is given by

$$\widehat{K}_h^{2h}(\theta^{00}) = \widehat{C}_h(\theta^{00})\widehat{S}_h^\nu(\theta^{00}) = (\widehat{I}_h(\theta^{00}) - \widehat{P}_{2h}^h(\theta^{00})\widehat{L}_{2h}^{-1}(\theta^{00})\widehat{R}_h^{2h}(\theta^{00})\widehat{L}_h(\theta^{00}))\widehat{S}_h^\nu(\theta^{00}).$$

Finally, we can compute the asymptotic two-grid convergence factor as the supremum of the spectral radii of $(n^2 \times n^2)$−matrices, as follows

$$\rho(K_h^{2h}) = \sup_{\theta^{00} \in \widetilde{\Theta}_{2nh}} \rho(\widehat{K}_h^{2h}(\theta^{00})),$$

where $\widetilde{\Theta}_{2nh}$ is the subset of Θ_{2nh} in which we remove the frequencies θ^{00} such that the determinant of the Fourier symbol of L_h or L_{2h} vanishes.

4.4. LFA results for PDEs with jumping coefficients

In this section, we apply the proposed LFA to predict the multigrid convergence factors for a collection of benchmark problems with jumping coefficients taken from the literature [63, 78, 83]. The test cases cover a variety of possible inhomogeneities including jumps that are not aligned with the coarse grid. In all these problems, equation (4.1) is numerically solved in the domain $\mathcal{D} = (0, 1)^2$, by using a mesh of grid-size $h = 1/128$. In particular, the following jumping coefficient benchmark problems, characterized by the distribution of the diffusion coefficient, are considered here:

1. **Vertical jump.** Function $k(x, y)$ is defined in the following way (see also Figure 4.3 (a))

$$k(x, y) = \begin{cases} 1, & \text{if } x < \frac{1}{2} + h, \\ 10^3, & \text{if } x \geq \frac{1}{2} + h. \end{cases}$$

2. **Four corner problem.** The domain is divided into four regions in which the diffusion coefficient is varying, see Figure 4.3 (b). In particular,

$$k(x, y) = \begin{cases} 10^4, & \text{if } (x, y) \in \left(0, \frac{1}{2} + h\right)^2, \\ 1, & \text{if } (x, y) \in \left(0, \frac{1}{2} + h\right) \times \left(\frac{1}{2} + h, 1\right), \\ 10^{-2} & \text{if } (x, y) \in \left(\frac{1}{2} + h, 1\right) \times \left(\frac{1}{2} + h, 1\right), \\ 10^{-4} & \text{otherwise.} \end{cases}$$

3. **Square inclusion.** In this example we assume a square inhomogeneity in one cell within the square domain, see Figure 4.3 (c). The diffusion coefficient is defined as

$$k(x, y) = \begin{cases} k_0, & \text{if } (x, y) \in \left(\frac{1}{2} - h, \frac{1}{2}\right)^2, \\ 1, & \text{otherwise,} \end{cases}$$

   where values $k_0 = 10^4$ and $k_0 = 10^{-4}$ are considered.

4. **Periodic square inclusions.** This test is taken from [64]. We consider a structured pattern of square inclusions of size $2h \times 2h$ as depicted in Figure 4.3 (d). The diffusion parameter is $k(x, y) = 1$ inside the dark region and $k(x, y) = 1000$ inside the white region.

5. **Periodic L-shaped inclusions.** In the last test case, we consider a structured pattern of L-shaped inclusions as in Figure 4.3 (e). The diffusion parameter is $k(x, y) = 10^4$ inside the white region and $k(x, y) = 1$ inside the dark region.
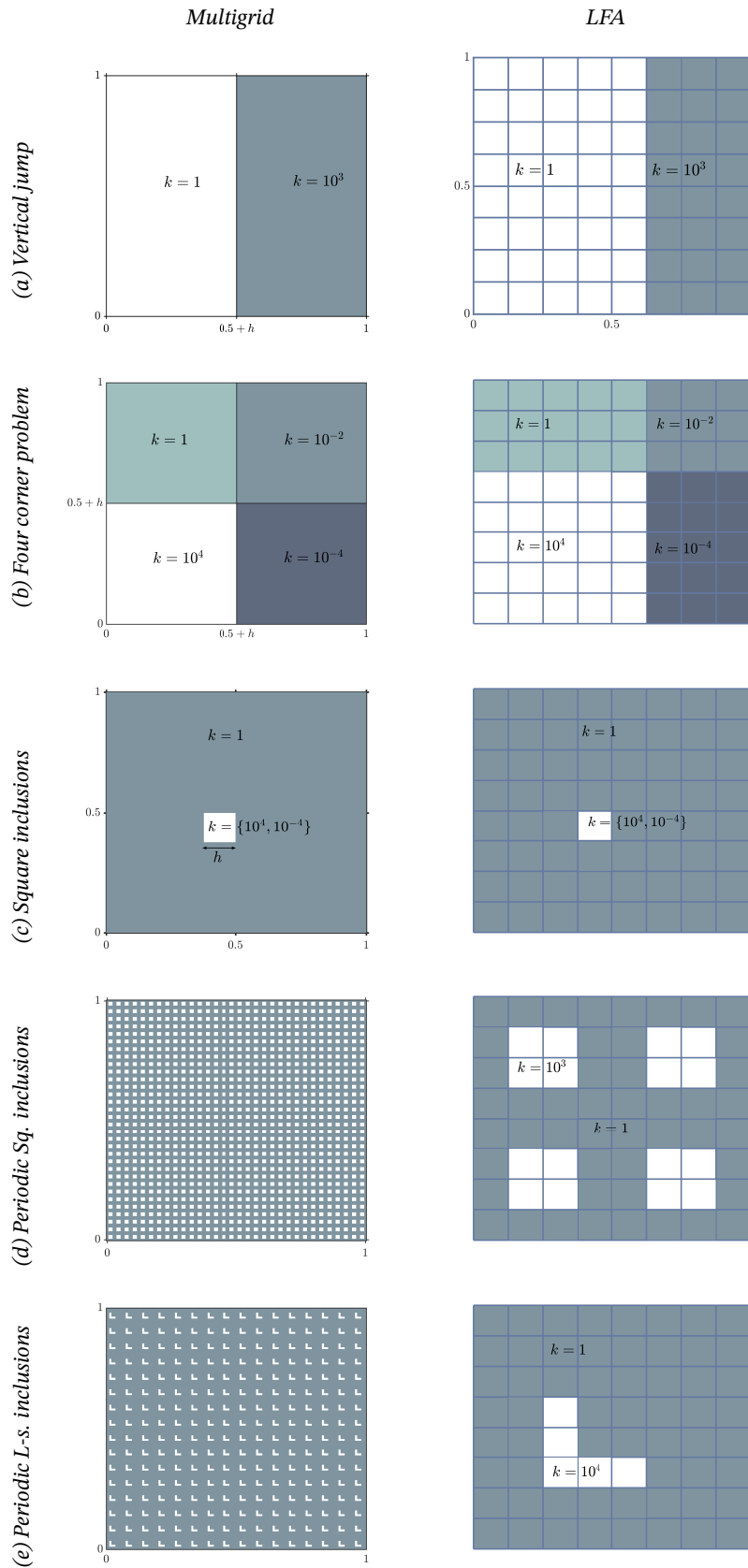
Figure 4.3: Distribution of the diffusion coefficients for the five considered examples on a unit square domain and corresponding 8×8 window used in the LFA.

To perform the theoretical analysis, the periodic extension of a window of size 8×8 has been chosen, where the diffusion coefficient is prescribed in such a window according to its definition, see Figure 4.3 (right side). In all numerical tests a random initial guess is chosen, and the right-hand side and boundary conditions are set to zero to be able to determine asymptotic convergence factors. In this way, we avoid round-off errors permitting us to perform as many iterations as needed. In practice, we have seen that 50 iterations are sufficient.

Next, we show the excellent correspondence between the theoretical analysis and the experimental results for these test cases. Two combinations of inter-grid transfer operators are considered. The first combination, denoted here by (CP,CR), is based on the use of piecewise constant prolongation (4.6) and its adjoint as the restriction (4.7). In the second combination we change to a higher polynomial order interpolation operator which is the adjoint to the Wesseling/Khalil restriction (4.21). This choice is denoted by (WP,CR). Moreover, a standard damped Jacobi smoother (damping with $\omega = 0.8$) is considered as well as the proposed lexicographic Gauss-Seidel smoother.

			Pre-, post-smoothing steps		
			$(1,0)$	$(1,1)$	$(2,2)$
Vertical jump	GS	(CP,CR)	0.42(0.42)	0.18(0.19)	0.04(0.03)
		(WP,CR)	0.41(0.37)	0.19(0.16)	0.07(0.11)
	Jac.	(CP,CR)	0.65(0.65)	0.43(0.42)	0.19(0.19)
		(WP,CR)	0.63(0.59)	0.40(0.35)	0.19(0.19)
Four corner problem	GS	(CP,CR)	0.42(0.37)	0.15(0.12)	0.04(0.03)
		(WP,CR)	0.40(0.39)	0.16(0.16)	0.09(0.09)
	Jac.	(CP,CR)	0.63(0.61)	0.40(0.39)	0.16(0.16)
		(WP,CR)	0.62(0.62)	0.40(0.40)	0.19(0.18)
Sq. inclusion ($k = 10^4$)	GS	(CP,CR)	0.45(0.44)	0.21(0.19)	0.04(0.04)
		(WP,CR)	0.41(0.40)	0.18(0.17)	0.11(0.11)
	Jac.	(CP,CR)	0.60(0.65)	0.36(0.42)	0.13(0.19)
		(WP,CR)	0.60(0.62)	0.38(0.37)	0.22(0.21)
Sq. inclusion ($k = 10^{-4}$)	GS	(CP,CR)	0.46(0.45)	0.21(0.20)	0.05(0.05)
		(WP,CR)	0.41(0.40)	0.19(0.19)	0.12(0.12)
	Jac.	(CP,CR)	0.61(0.65)	0.38(0.42)	0.15(0.19)
		(WP,CR)	0.61(0.59)	0.39(0.39)	0.23(0.23)
Periodic Sq. inclusions	GS	(CP,CR)	0.64(0.61)	0.43(0.42)	0.41(0.41)
		(WP,CR)	0.62(0.61)	0.43(0.41)	0.41(0.40)
	Jac.	(CP,CR)	0.81(0.78)	0.66(0.65)	0.44(0.46)
		(WP,CR)	0.81(0.80)	0.66(0.65)	0.44(0.43)
Periodic L-S. inclusions	GS	(CP,CR)	0.50(0.50)	0.32(0.26)	0.21(0.21)
		(WP,CR)	0.54(0.53)	0.40(0.40)	0.30(0.30)
	Jac.	(CP,CR)	0.71(0.63)	0.54(0.48)	0.36(0.35)
		(WP,CR)	0.71(0.71)	0.56(0.56)	0.42(0.42)

Table 4.1: Asymptotic two-grid convergence factors predicted by LFA and the corresponding computed average multigrid convergence factors (in parenthesis) using two-grid cycles and different pre- and post-smoothing steps for the five examples.

In Table 4.1, for different numbers of smoothing steps, for two different smoothers and for the two combinations of restriction and prolongation operators, we provide the two-grid convergence factors predicted by the novel LFA for each of the proposed numerical experiments. We also display in parenthesis the average after 50 iterations of the experimentally computed multigrid convergence factors by using two-grid cycles. For all these cases, we observe a very accurate match between the

analysis results and the rates experimentally obtained. Regarding the size of the window to perform the LFA, we have observed that a window of size 8×8 is enough to achieve excellent predictions in all considered benchmark problems. For example, for the vertical jump test, the two-grid analysis considering four smoothing steps of the Gauss-Seidel smoother and the combination (CP,CR) of inter-grid transfer operators provides a factor of 0.11 when a 2×2 window is used, a factor of 0.06 for a 4×4 window, and a factor of 0.04, which matches perfectly the real convergence, when the 8×8 window is considered.

We remark that with the current multigrid approach, the quality of the coarse grid discretization may not be satisfactory, for example, for chessboard or L-shaped inclusion examples. For such cases, we either recommend using more *powerful* smoothers such as the ILU smoother or adapt the coarsening. Furthermore, for PDEs with strong local variations in the coefficient fields, *homogenization techniques* [63, 78, 83] may also be used to obtain the coarser representation of the fine grid problem. Comparing the results of the two combinations of inter-grid transfer operators, we observe a very similar performance for all five test cases studied here. In the rest of the chapter, we, therefore, choose the strategy (CP,CR) because of its simplicity and low computational cost.

4.5. LFA results for PDEs with random coefficients

Here, we consider the SPDE (4.5) defined on a unit square domain $\mathcal{D} = (0, 1)^2$ with homogeneous Dirichlet boundary conditions. Two different types of diffusion coefficients based on random jumps and lognormal random fields are studied. The randomly jumping coefficient problem can be seen as a transition from the deterministic to a stochastic setting.

4.5.1 Randomly jumping coefficients

To simulate random jumps, the domain $\mathcal{D}$ is subdivided into square-blocks of size $[1/8 \times 1/8]$ and the value of the coefficients on each of the blocks is sampled as

$$k = e^U \quad \text{with} \quad U \sim \mathcal{U}\{-m, m\} \quad \text{and} \quad m \in \mathbb{Z}. \tag{4.22}$$

In other words, U is an independent identically distributed (i.i.d.) integer sampled from a discrete uniform distribution $\mathcal{U}\{-m, m\}$. Here, the integer m defines the order of magnitude of the jumps, an example for $m = 5$ is shown in Figure 4.4. Notice that for this choice of m, we may encounter interfaces with maximum jumps of magnitude equal to e^{10}. For each random realization of the jumping coefficient field, we compare the LFA two-grid convergence factors with the computed asymptotic convergence factors of the multigrid method by using W-cycles. To perform the LFA, we again use a window of size 8×8. Furthermore, regarding LFA the randomly jumping coefficient problem is similar to the four corner problem in Figure 4.3 (d) with the magnitude of each block given by (4.22) and the cross-point exactly at the center of the LFA block. Regarding the multigrid components, a lexicographic Gauss-Seidel iteration is employed as the smoother, and the simplest combination (CP,CR) of inter-grid transfer operators is chosen. Also, we use a 4×4 grid as the coarsest in the multigrid hierarchy. To determine the asymptotic convergence factors of the multigrid method, the right-hand side is again set to zero.

The experimental convergence factor of the multigrid method for the i-th realization of the random field is then computed, as follows:

$$\rho_i = \left\{ \frac{||res^{\kappa_i}||_\infty}{||res^0||_\infty} \right\}^{1/\kappa_i}, \qquad \text{for} \quad i = 1, 2, ..., N_{MG}, \tag{4.23}$$

where $||res^0||_\infty$ is the infinity norm of the residual obtained from an initial solution and $||res^{\kappa_i}||_\infty$ is the residual after κ_i iterations of the multigrid cycle. We use these quantities to calculate the
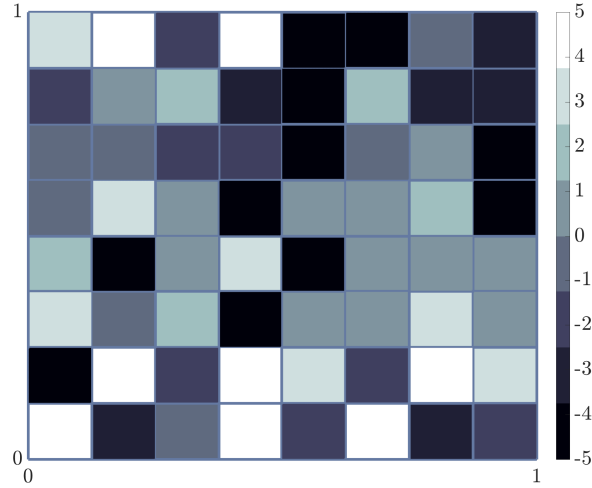
Figure 4.4: An example of random realization of $U = \ln k$ with $m = 5$ on a unit square domain.

average and the standard deviation of the asymptotic convergence factors:

$$\langle \rho \rangle_{MG} = \frac{1}{N_{MG}} \sum_{i=1}^{N_{MG}} \rho_i, \qquad \sigma_{MG} = \sqrt{\frac{1}{(N_{MG} - 1)} \sum_{i=1}^{N_{MG}} (\rho_i - \langle \rho \rangle_{MG})^2}, \qquad (4.24)$$

respectively. These averaged quantities are defined similarly for the LFA results (based on LFA two-grid factors), and are denoted, respectively, by $\langle \rho \rangle_{LFA}$ and σ_{LFA}.

In Figure 4.5, we show the comparison, the mean $\pm$ standard deviation, of the LFA prediction and the multigrid convergence for jump parameter $m = 2$ (left) and $m = 5$ (right) computed using $N_{LFA} = N_{MG} = 100$. Overall, a good match between the LFA and MG convergence is seen up to one decimal place. We also observe that for this specific jumping coefficient problem, there is no further improvement with an increase in the number of smoothing steps after the $W(2,2)$−cycle.
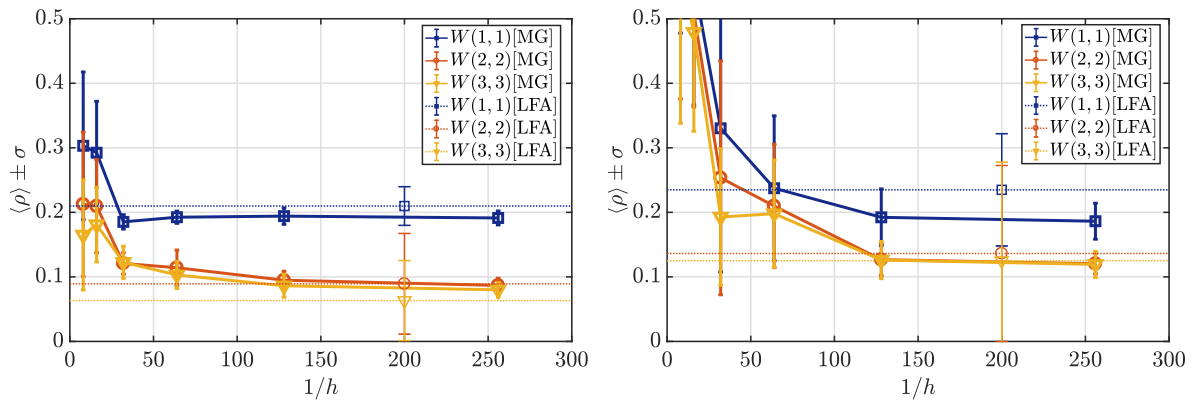

Figure 4.5: Comparison of the mean and the standard deviation of the LFA (dashed line) and MG (solid line) convergence factors for different W−cycling strategies for randomly jumping coefficients with $m = 2$ (left) and $m = 5$ (right).

4.5.2 Lognormal random fields

Next, we test the LFA prediction capability for elliptic PDEs with lognormal diffusion coefficients. We consider four Matérn reference parameter sets Φ with increasing order of complexity, listed in Table 4.2. Examples of random fields generated with these Matérn parameter sets were already presented in Section 3.2. For a fixed Φ and mesh size h, we generate N_{MG} realizations of the

Table 4.2: Different combinations of the Matérn reference parameters $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$ with increasing complexity from left to right.

Φ_1	Φ_2	Φ_3	Φ_4
(1.5,0.3,1)	(0.5,0.3,1)	(1.5,0.1,3)	(0.5,0.1,3)

permeability field. For each of these random fields, we can compare the LFA two-grid convergence factors with the computed asymptotic convergence factors of the multigrid method by using the same components that were utilized for randomly jumping coefficient fields. The comparison is shown in Table 4.3, where different numbers of smoothing steps are considered and the comparison is done for the four reference parameter sets, described in Table 4.2. For all experiments, we set $h = 1/64$ and $N_{MG} = N_{LFA} = 100$. In general, we observe a good agreement between the experimental and the LFA quantities. For the first three sets of parameters excellent convergence factors are obtained already with two pre- and two post-smoothing steps. When the more difficult set of parameters is considered, however, more smoothing steps may be necessary to obtain a good convergence factor. It is also pointed out that the performance of W(1,1)-cycle is well predicted by LFA for all Φ_i, $i = 1, 4$. A slight discrepancy is observed for the W(2,2)- and W(3,3)-cycles in the case of Φ_4.

		Φ_1		Φ_2		Φ_3		Φ_4	
		MG	LFA	MG	LFA	MG	LFA	MG	LFA
$W(1,1)$	$\langle\rho\rangle$	0.20	0.20	0.20	0.20	0.19	0.20	0.23	0.21
	σ	0.004	0.002	0.004	0.005	0.005	0.01	0.04	0.02
$W(2,2)$	$\langle\rho\rangle$	0.04	0.04	0.04	0.04	0.07	0.06	0.17	0.08
	σ	0.001	0.001	0.002	0.003	0.03	0.03	0.04	0.03
$W(3,3)$	$\langle\rho\rangle$	0.01	0.02	0.02	0.02	0.05	0.03	0.13	0.06
	σ	0.001	0.002	0.002	0.004	0.02	0.02	0.03	0.03

Table 4.3: Comparison of the average and the standard deviation of the LFA and MG convergence factors. Different numbers of smoothing steps, $h = 1/64$, and the four reference parameter sets in Table 4.2 are considered.

4.5.3 Mesh dependency

Next, we wish to study the influence of the size of the sampling mesh on the multigrid convergence. For this purpose, we choose two representative parameter sets describing a smooth and a highly oscillating random field, that are generated by the parameter sets Φ_2 and Φ_4, respectively. Figure 4.6 shows the average convergence factors for Φ_2 (left side) and Φ_4 (right side), predicted by LFA (top) and experimentally observed multigrid convergence (bottom), for different mesh sizes and different numbers of smoothing steps. For Φ_2, the average reduction factor is roughly the same, independently of the size of the mesh. These predictions coincide well with the experimentally observed factors. For the results corresponding to parameter set Φ_4, we observe robustness of the

method when the mesh is sufficiently fine, which is also confirmed by the multigrid experiments. In the same figure, the standard deviation is presented, which decreases when $h \to 0$. The LFA predictions on the coarser grids are less reliable when compared to the multigrid convergence rates. We would like to mention that a three-grid LFA [84] is not helpful here. On coarse grids, boundary conditions have an impact on the method's convergence, but they are not taken into the account in the analysis. Moreover, to analyze the robustness of the proposed multigrid method,
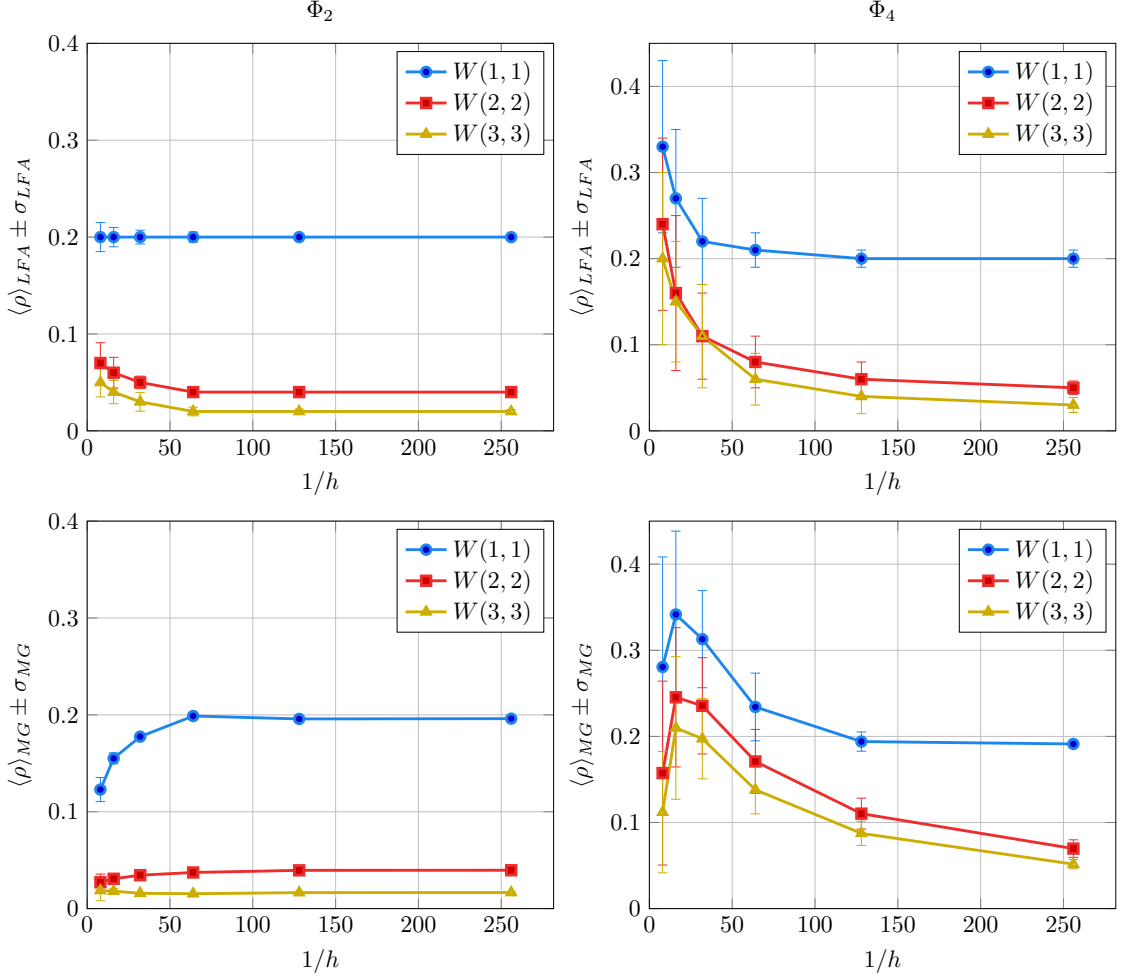

Figure 4.6: Average reduction LFA two-grid factors (top) and asymptotic MG convergence (bottom) along with the standard deviation for two reference parameter sets Φ_2 (left column) and Φ_4 (right column); $N_{MG} = N_{LFA} = 100$.

in the next experiment we fix the field's smoothness parameter as $\nu_c = 0.5$, and vary the other parameters $\sigma_c^2 \in [0.5, 5]$ and $\lambda_c \in [0.05, 0.5]$. In Figure 4.7 we show the average LFA two-grid convergence factors when two smoothing steps are considered for $h = 1/32$ (left) and $h = 1/64$ (right). The multigrid convergence is very satisfactory for all combinations of the parameters and increases slightly when λ_c tends to be small and σ_c^2 becomes large. This case, however, represents a rather extreme situation in which the jumps in the permeability field are of more than 15 orders of magnitude. Again, we see an improvement in the convergence rate with grid refinement. Note that with the considered range of parameters we cover all realistic cases.
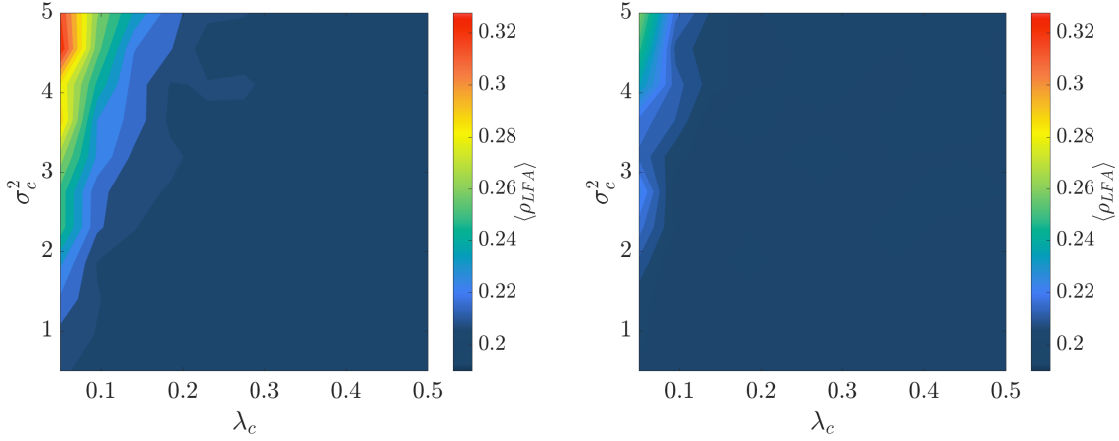
Figure 4.7: Contour of average LFA two-grid convergence factors, $\langle \rho \rangle_{LFA}$, with two smoothing steps for different λ_c and σ_c^2, and for fixed $\nu_c = 0.5$ with $h = 1/32$ (left) and $h = 1/64$ (right).

4.6. Multigrid performance for anisotropic random fields

We study the performance of the cell-centered multigrid for anisotropic random fields. To generate these random fields, we use a modified version of the Matérn covariance function given by

$$C_{\widetilde{\Phi}}(\mathbf{x_1}, \mathbf{x_2}) = \sigma_c^2 \frac{2^{1-\nu_c}}{\Gamma(\nu_c)} \left(2\sqrt{\nu_c} \tilde{r} \right)^{\nu_c} K_{\nu_c} \left(2\sqrt{\nu_c} \tilde{r} \right) \quad \mathbf{x_1}, \mathbf{x_2} \in \mathcal{D}, \tag{4.25}$$

where $C_{\widetilde{\Phi}}$ is a stationary covariance function depending on the parameter set $\widetilde{\Phi} = (\nu_c, \lambda_{cx}, \lambda_{cy}, \sigma_c^2, \theta)$. The distance function $\tilde{r}$ is defined as

$$\tilde{r} = \sqrt{\frac{(x_1' - x_2')^2}{\lambda_{cx}^2} + \frac{(y_1' - y_2')^2}{\lambda_{cy}^2}}, \tag{4.26}$$

$$x_1' = x_1 \cos\theta - y_1 \sin\theta,$$

$$y_1' = x_1 \sin\theta + y_1 \cos\theta, \quad \text{with} \quad \mathbf{x_1} = (x_1, y_1).$$

Here (x_1', y_1') is obtained by rotating (x_1, y_1) by angle θ in counterclockwise direction with respect to the horizontal axis. The quantities λ_{cx} and λ_{cy} are correlation lengths along the x- and y-coordinates, respectively. The covariance function $C_{\widetilde{\Phi}}$ only differs from the isotropic covariance C_{Φ} defined in (3.7) in terms of the distance function $\tilde{r}$. In Figure 4.8, we present realizations of the anisotropic random field (generated using circulant embedding, see Section 3.6) with two different $\widetilde{\Phi}$ values. As the two parameter sets only differ in terms of the rotation parameter θ, they exhibit a similar magnitude of the jumps. Note that the random fields generated from $\widetilde{\Phi}_2$ are more challenging for the cell-centered multigrid as the long-range correlations are not aligned with the grid.

In Figure 4.9, we show the mean $\pm$ standard deviation of the asymptotic multigrid convergence for the two parameter sets, $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$. Although, we see some deterioration compared to the isotropic case, the convergence rates improve with grid refinement. Here, using the x-line smoother for horizontal layering and the alternating line smoother for the non-aligned random field, keeping the other components same, will further improve the convergence.
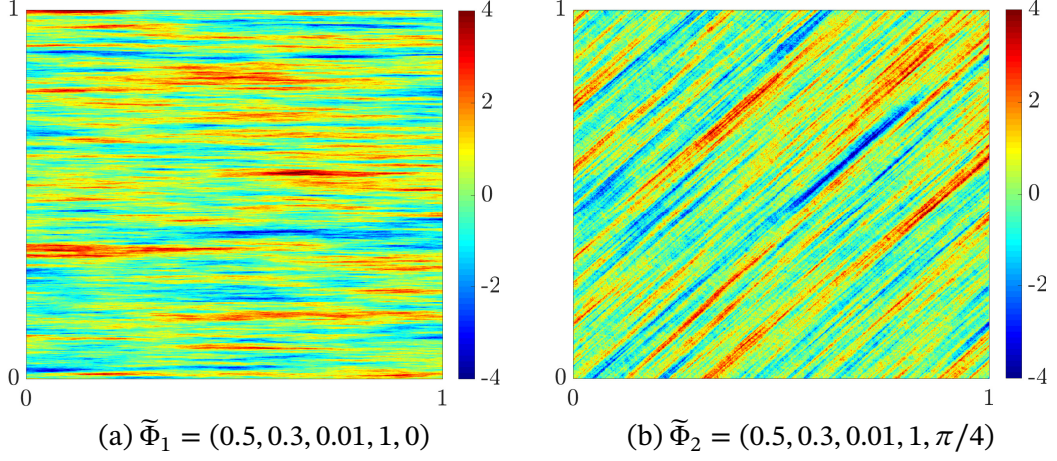
(a) $\widetilde{\Phi}_1 = (0.5, 0.3, 0.01, 1, 0)$ (b) $\widetilde{\Phi}_2 = (0.5, 0.3, 0.01, 1, \pi/4)$

Figure 4.8: Logarithm of the permeability field, $\log_{10} k$, generated using two reference parameter sets on a unit square domain.
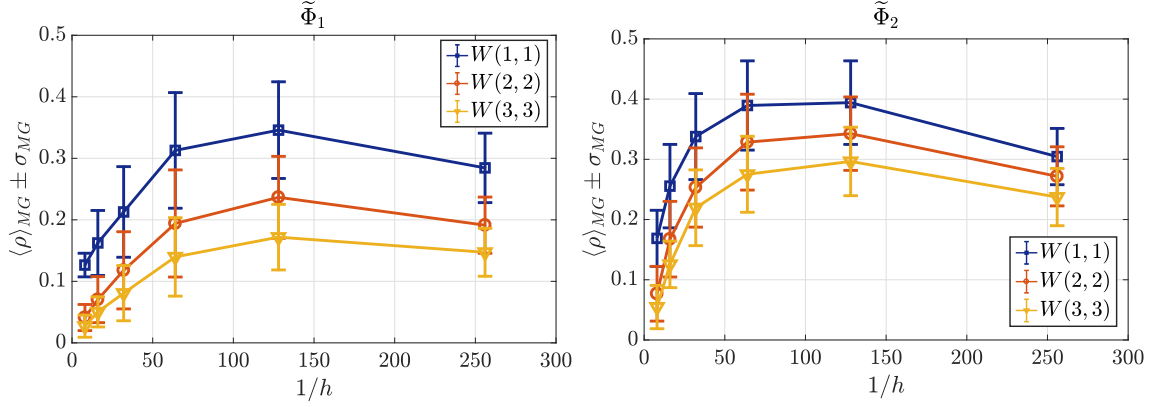

Figure 4.9: Mean and the standard deviation of the asymptotic MG convergence for two reference parameter sets $\widetilde{\Phi}_1$ (left) and $\widetilde{\Phi}_2$ (right) using $N_{MG} = 100$.

4.7. MLMC numerical experiments

In this section we analyze the performance of the proposed multigrid MLMC method. We consider PDE (4.5) on domain $\mathcal{D} \in (0,1)^2$ with mixed Dirichlet-Neumann boundary conditions,

$$p(0, y, \omega) = 1, \qquad p(1, y, \omega) = 0, \quad \text{and} \tag{4.27}$$
$$k(\mathbf{x}, \omega)\frac{\partial p}{\partial x}\bigg|_{y=0} = 0, \qquad k(\mathbf{x}, \omega)\frac{\partial p}{\partial x}\bigg|_{y=1} = 0,$$

respectively. For all tests the source term is set to zero, i.e. $f = 0$. As the quantity of interest, the outflow through the boundary $x = 1, y = (0,1)$, also referred to as *effective permeability*, is considered:

$$Q(p) = -\int_0^1 k(\mathbf{x}, \omega)\frac{\partial p}{\partial x}(\mathbf{x}, \omega)\bigg|_{x=1} dy. \tag{4.28}$$

Our aim is to compute $\mathbb{E}[Q]$ using the MLMC estimator outlined in Chapter 2. In Figure 4.10, we show the convergence of the FV bias, $|\mathbb{E}[Q_{h_\ell} - Q_{h_{\ell-1}}]|$, and the level dependent variance

$\mathcal{V}_\ell = \mathbb{V}[Q_{h_\ell} - Q_{h_{\ell-1}}]$ for the considered quantity of interest (4.28). We observe that the rate of decay of the FV bias depends on the smoothness parameter ν_c of the random fields. In the case of parameter sets Φ_1 and Φ_3, we see a second-order convergence whereas a first-order convergence is observed for Φ_2 and Φ_4. The correlation length λ_c and variance σ_c^2 only affect the proportionality constant. Also, for all four cases, the level dependent variance shows a quadratic decay. Next,
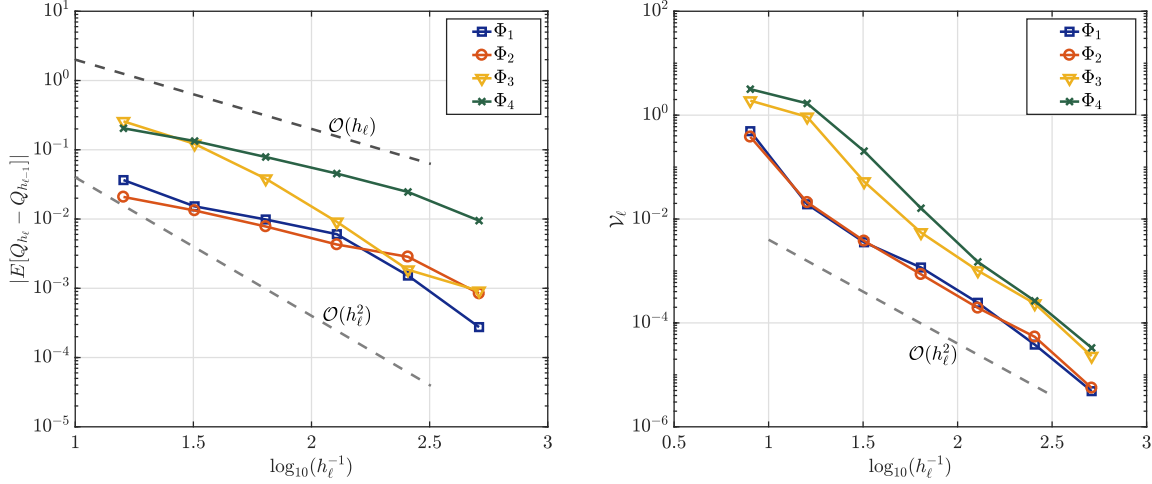

Figure 4.10: (Left) Convergence of finite volume error, (Right) decay of level dependent variance for the outflow, Q_{h_ℓ}, with mesh refinement for different Matérn parameters.

we conduct MLMC experiments for Matérn parameter sets Φ_2 and Φ_4, using W(1,1), W(2,2) and W(3,3)-cycles. The multigrid components described in Section 4.2.2 are used with stopping criterion $\varepsilon_{MG} = 10^{-10}$. We consider $\alpha = 1$ and $\beta = 2$ for both parameter sets Φ_2 and Φ_4. For all experiments, we use the fixed sampling strategy described in Section 2.3.3. This sampling strategy requires the total number of MLMC levels, the number of finest level samples N_L along with the rate β. We set $N_L = 64$ within MLMC for both parameter sets resulting in a sample sequence $N_\ell = N_L 4^{(L-\ell)}$.

4.7.1   MLMC results for Φ_2

We first analyze the error convergence of the expected value and the variance of the quantity of interest. In Figure 4.11 (left), relative errors with respect to a reference solution computed using the MLMC estimator with the finest level $h^* = 1/512$ are shown. We see the error $|\mathcal{E}_L^{ML}[Q_{h_L}] - \mathcal{E}_{ref}^{ML}[Q_{h^*}]|$ converges with $\mathcal{O}(h_L)$. Next, we analyze the performance of the MLMC estimator for different multigrid cycling strategies. In Figure 4.11 (right), we present the CPU-times for the MLMC simulations with an increasing number of MLMC levels. These CPU-times are derived by summing up the run-times from the multigrid solves for all samples over all levels. Here, we do not include the cost of generation of the random field and the post-processing costs as they are the same for different cycles. The cost scales as $\mathcal{O}(h_L^{-2})$ coinciding with the theoretical MLMC complexity when $\beta = \gamma$. No substantial difference in the CPU-times from the three cycles is observed. We also show the level-wise CPU-times in Table 4.4 for the three cycles with the same number of MLMC samples per level. In general, the computational cost improves with levels for all cycles. We observe that the W(2,2)-cycle slightly outperforms the other two variants.

In Figure 4.12, we also show the distribution of the number of multigrid iterations for the W(2,2)-cycle on different grid levels along with the number of iterations predicted by the LFA, $\tilde{\kappa}_\ell := \lceil \log(\varepsilon_{MG}) / \log(\langle\rho\rangle_{LFA,\ell}) \rceil$. The cost is more heterogeneous on coarser levels with larger variance in the number of iterations and is homogeneous for $h_\ell = 1/32$ and onwards. A high

variance in the multigrid convergence rate was also predicted by the LFA experiments in the previous section. Thus, a sparse direct solver can replace the multigrid solver on these coarser levels. In this work, the considered multigrid stopping criteria of 10^{-10} is quite conservative. For many engineering applications a residual reduction of 10^{-6} may already be sufficient to reach a converged solution, therefore, reducing the computational cost roughly by a factor of two.
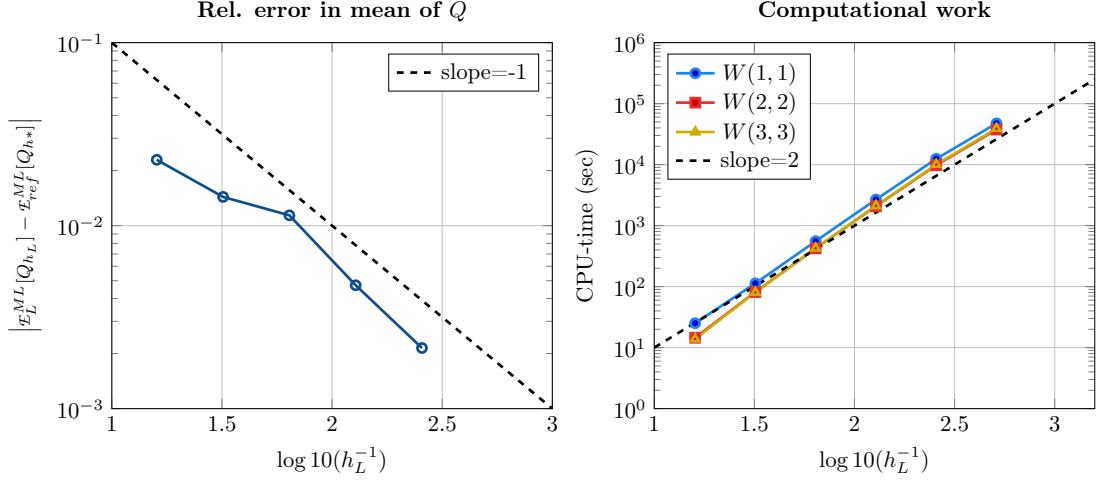

Figure 4.11: (Left) Convergence of error in $\mathcal{E}_L^{ML}[Q_{h_L}]$ with increasing number of MLMC levels for parameter Φ_2. The reference solution is based on mesh size $h^* = 1/512$. (Right) Mean CPU-times versus accuracy for different W-cycles. Computational cost proportional to $\mathcal{O}(h_L^{-2})$ is observed for all cycling strategies.

		level-wise CPU-time (sec.)		
h_ℓ	N_ℓ	W(1,1)	W(2,2)	W(3,3)
1/8	262144	489	331	323
1/16	65536	528	391	390
1/32	16384	454	362	363
1/64	4096	407	336	348
1/128	1024	376	320	342
1/256	256	357	308	338
1/512	64	350	300	333

Table 4.4: Comparison of the three W-cycles in terms of the level-wise CPU-times for parameter Φ_2.

4.7.2 MLMC results for Φ_4

Now we describe MLMC results for the challenging parameter set Φ_4 for the same quantity of interest. In Figure 4.13 (left), we display the convergence of the relative error in the expected value of the quantity of interest using a reference solution obtained with finest resolution $h^* = 1/1024$. Due to a large variance σ_c and a small correlation length λ_c of the random field, the error is larger, as compared to the results of parameter set Φ_2 for same mesh size h_L and the *expected* convergence rates are visible on relatively fine grids. A similar trend is observed for the computational cost with scaling as $\mathcal{O}(h_L^{-2})$ in Figure 4.13 (right). Again, in Table 4.5, the level-wise CPU-times for the three cycles are listed. Similar to the results for set Φ_2, the cost improves with levels but the gain from W(2,2)-cycles is more prominent.
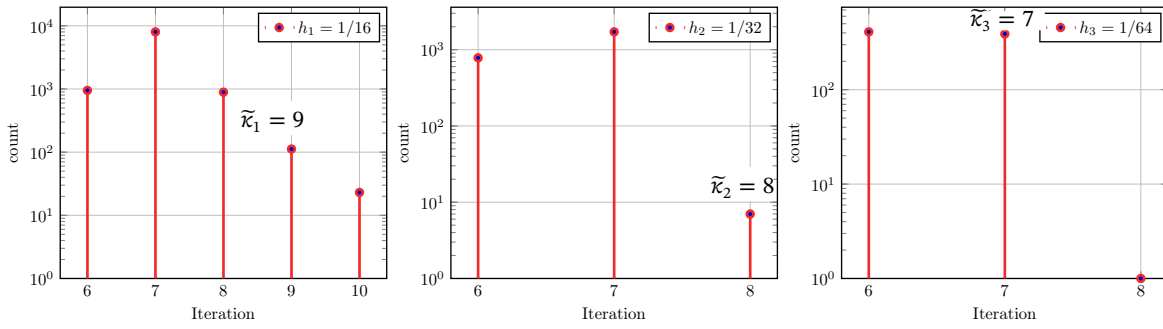
Figure 4.12: Distribution of the number of multigrid iterations of $W(2,2)$-cycles for different mesh sizes for parameter Φ_2 to reach $\varepsilon_{MG} < 10^{-10}$ and the LFA predicted number of iteration $\widetilde{\kappa}_\ell$.

Lastly, the distribution of the number of iterations for W(2,2)-cycles is depicted in Figure 4.14. High variability in the number of multigrid iterations persists until a relatively fine grid ($h = 1/128$). However, the average number of iterations predicted by LFA coincides very well with the mode of these distributions. Again, a direct solver can be applied on these coarser levels to get a more reliable estimate of the cost. Here, we point out that the W(2,2)-cycle takes roughly 7 iterations similar to parameter Φ_2 to reduce the residual by 10^{-10}, indicating robustness with respect to the Matérn parameter.
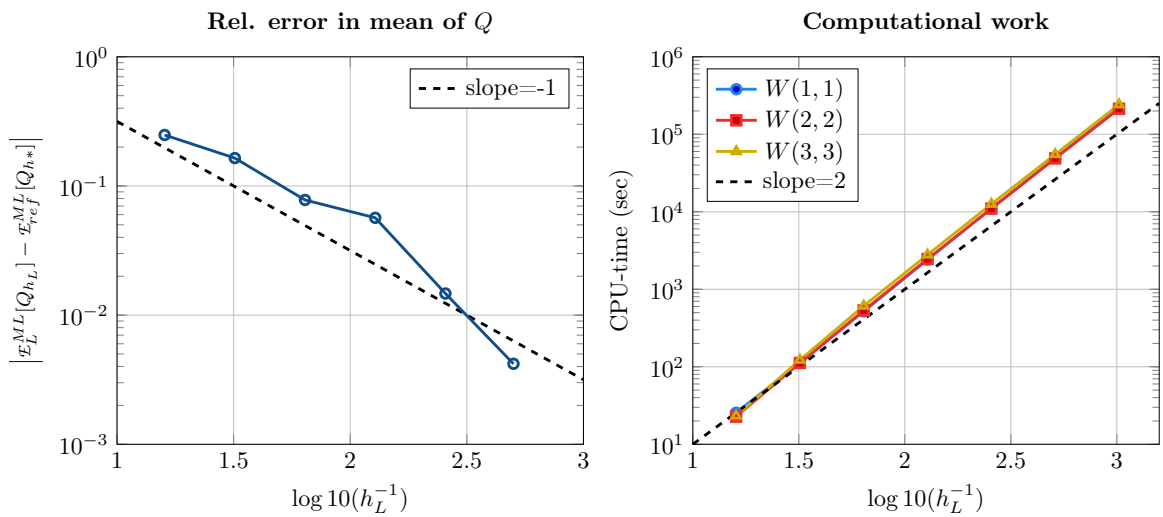

Figure 4.13: (Left) Convergence of error in $\mathcal{E}_L^{ML}[Q_{h_L}]$ with increasing number of MLMC levels for parameter Φ_4. The reference solution is based on mesh size $h^* = 1/1024$. (Right) Mean CPU-times versus accuracy for different W-cycles. Computational cost proportional to $\mathcal{O}(h_L^{-2})$ is observed for all cycling strategies.

		level-wise CPU-time (sec.)		
h_ℓ	N_ℓ	W(1,1)	W(2,2)	W(3,3)
1/8	1048576	2.23[+3]	2.10[+3]	2.18[+2]
1/16	262144	2.42[+3]	2.52[+3]	2.80[+3]
1/32	65536	1.85[+3]	2.10[+3]	2.45[+3]
1/64	16384	1.56[+3]	1.59[+3]	1.98[+3]
1/128	4096	1.45[+3]	1.33[+3]	1.60[+3]
1/256	1024	1.40[+3]	1.23[+3]	1.40[+3]
1/512	256	1.38[+3]	1.23[+3]	1.34[+3]
1/1024	64	1.36[+3]	1.21[+3]	1.31[+3]

Table 4.5: Comparison of the three W-cycles in terms of the level-wise CPU-times for parameter Φ_4.
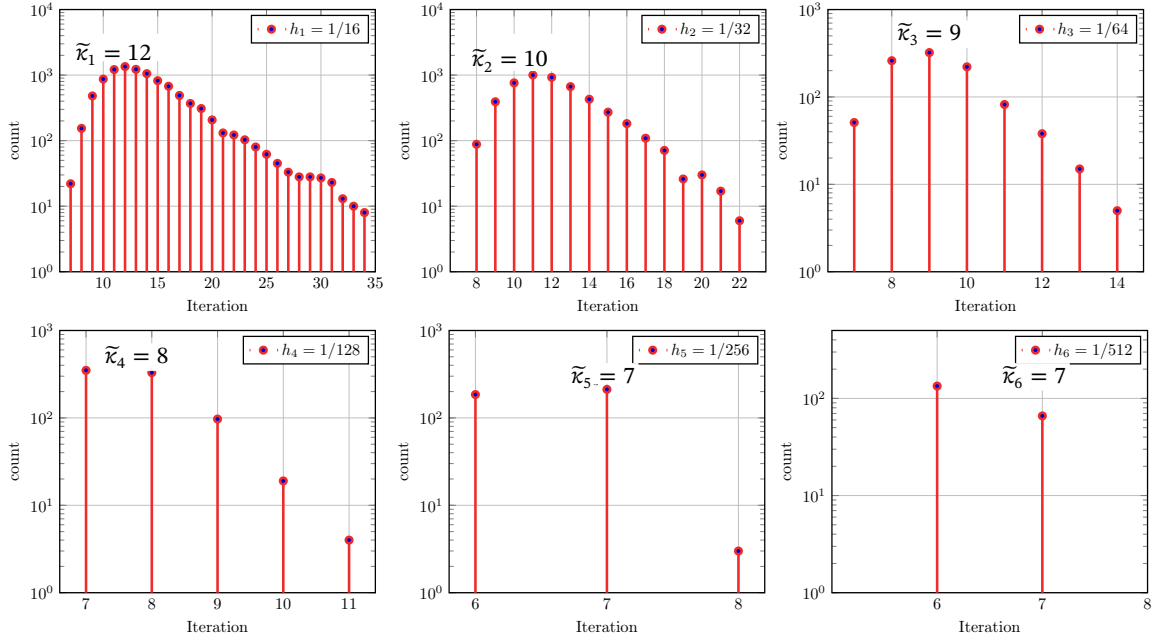

Figure 4.14: Distribution of the number of multigrid iterations of $W(2,2)$-cycles for different mesh sizes for parameter Φ_4 to reach $\varepsilon_{MG} < 10^{-10}$ and the LFA predicted number of iteration $\widetilde{\kappa}_\ell$.

4.8. Conclusions

A novel, generalized Local Fourier Analysis which can be employed for quantitative assessment of multigrid methods for PDEs involving jumping and random coefficients has been presented. In particular, a cell-centered multigrid algorithm for solving a model problem based on a single-phase flow problem in a random porous medium has been used to show the accuracy of the proposed analysis. This technique, however, is appropriate for the prediction of the performance of a wider class of multigrid methods for solving PDEs with random fields. The effectiveness of the proposed LFA method was also confirmed numerically using a number of challenging test cases with jumping and random coefficients. Further, the novel local Fourier analysis can help us to estimate apriori the time needed for solving certain uncertainty quantification problems by using a multigrid multilevel Monte Carlo method.

<div align="right">

5

</div>

Transport in a coupled Darcy-Stokes system

In this chapter, we concentrate on the uncertainty quantification of advection-dominated contaminant transport in a coupled Darcy-Stokes flow system using the multigrid multilevel Monte Carlo method. The uncertainty in the flow arises due to an unknown permeability of the porous media that is modeled as a lognormal random field. Different numerical strategies for the subproblems are explored and an optimal combination for the MLMC estimator is then suggested. A specific monolithic multigrid algorithm is also presented to efficiently solve the steady-state Darcy-Stokes flow with a highly heterogeneous diffusion coefficient. The advection-dominated contaminant transport equation is numerically solved using the Alternating Direction Implicit (ADI) based time-stepping for the flux-limited quadratic upwinding discretization. In the final part of the chapter, some numerical experiments illustrating the multigrid convergence and cost of the MLMC estimator with respect to the smoothness of the permeability field are presented.

5.1. Introduction

Transport in a Darcy-Stokes flow system can be used to analyze a large number of dynamical processes. This model is, for example, of importance in the study of accidental discharge of radioactive contaminants or chemical spillage in surface water bodies and the subsequent transport to the connected aquifers. A coupled Darcy-Stokes system is then used to simulate the interaction between the surface water and the groundwater flow. The coupling is achieved by imposing interface conditions based on mass conservation, balancing the normal stress and a special condition, called the Beavers-Joseph-Saffman (BJS) interface condition [85, 86] that relates the shear stress and tangential velocity along the interface. The steady-state velocity field derived from this model is then utilized for the advection of chemical components in a transport model. Many relevant physical phenomena concerning mass transport such as molecular diffusion, mechanical dispersion, adsorption, can be conveniently incorporated in this model.

The mathematical theory and analysis of transport in a coupled Darcy-Stokes flow is well developed and a number of stable, convergent numerical methods have been proposed. These include finite volume (FV), mixed finite element (MFE) or more advanced, locally conservative discontinuous Galerkin (DG) schemes, see [87–94] and references therein. These schemes require

the knowledge of physical quantities like fluid viscosity, permeabilities and experimentally measured BJS interface parameters for approximating the Darcy-Stokes flow. Furthermore, the chemical transport equation also requires input such as initial and inflow boundary conditions. In many cases, complete information of these physical quantities is not available and they may be modeled in a probabilistic framework. For instance, it is well known that the permeability field can be modeled as a lognormal random field [18–20]. Once these uncertainties are incorporated in the mathematical models, the goal is to obtain the statistics of certain quantities of interest, for example, the mean spatial concentration of contaminants in an aquifer after a certain interval from the time of discharge.

The purpose of this chapter is to describe in a systematic manner a numerical strategy to design an efficient multilevel Monte Carlo estimator for UQ of stochastic transport in the Darcy-Stokes system. An efficient MLMC estimator requires careful consideration of the numerical techniques for the approximation of the QoI. The chapter is organized in the following way:

- In Section 5.2, we describe the stochastic transport in the Darcy-Stokes flow. The mixed formulation of the Darcy equation is used that is coupled with the Stokes flow using three interface conditions. The stochastic extension of the problem is obtained by modeling the permeability as a lognormal random field. The spatial covariance of the random field is derived from the parameterized Matérn function defined in (3.7).

- In Section 5.3, the Finite Volume (FV) discretization of the Darcy-Stokes flow on a *staggered grid* is described. A special discretization along the interface is proposed taking into account the Beaver-Joseph-Saffmann interface condition that depends on the random permeability along the interface and a specific parameter. The optimal choice for the spatio-temporal discretization of the transport equation is based on the regularity of the Darcy-Stokes solution. Typically, the error in the FV approximation of the velocity field depends on the spatial regularity of the permeability field and also on the smoothness of boundary conditions along the two domains. Thus, the discretization scheme for the transport equation should be related to the accuracy of the velocity approximation. Using higher-order schemes for low regularity problems may lead to an expensive MLMC estimator without any improvement in the numerical accuracy. We also show that, for advection-dominated transport with sharp gradients and discontinuities, low-order schemes are very diffusive and are less suited for MLMC applications.

- A monolithic multigrid solver for the Darcy-Stokes problem based on an Uzawa smoother [95] is proposed in Section 5.4. This smoother employs an equation-wise decoupled relaxation for the pressure and velocity unknowns. For the velocity, a symmetric Gauss-Seidel iteration is employed whereas for the pressure a Richardson iteration is applied. The Richardson iteration takes into account local fluctuations in the permeability field to derive the optimal relaxation parameter. The multigrid algorithm is based on cell-centered permeability coefficients and direct coarse-grid discretization based on cell-centered averaging. The proposed solver is robust and also works well on very coarse grids. We generalize this multigrid method to multi-block problems by the *grid partitioning technique* [14]. To incorporate the coarse time-step in the MLMC hierarchy, an implicit time stepping is desired for the contaminant transport equation. We consider an Alternating Direction Implicit (ADI) based solver for the discrete transport equation which breaks the problem into two 1D problems, greatly reducing the cost of time-stepping.

- Section 5.5 is devoted to numerical experiments performed on two test problems: a 2-block problem with no-slip interface condition and a more realistic 4-block case with the BJS interface condition. We thoroughly test the multigrid method with respect to different Matérn

parameters to identify the optimal cycling strategy. Finally, we combine these components to obtain the solution of the stochastic transport problem using MLMC estimator and compare the asymptotic cost with the standard Monte Carlo method.

5.2. Stochastic transport in Darcy-Stokes system

We consider the transport equation defined in the bounded domain $\mathcal{D} \subset \mathbb{R}^2$, with boundary $\partial \mathcal{D}$ and in a finite time interval $\mathcal{T} = (0, T]$, for $T < \infty$. The transport equation coupled with the steady-state Darcy-Stokes flow on $\mathcal{D}$, is subdivided into a porous medium $\mathcal{D}^d \subset \mathbb{R}^2$, where the flow is described by Darcy's law and the free-flow region $\mathcal{D}^s \subset \mathbb{R}^2$, governed by the Stokes equations with boundaries $\partial \mathcal{D}^d$ and $\partial \mathcal{D}^s$, respectively. The internal interface is defined as $\Gamma_{ds} = \partial \mathcal{D}^d \bigcap \partial \mathcal{D}^s$. Furthermore, we denote by ω an event in the probability space $(\Omega, \mathbb{F}, \mathbb{P})$, where Ω is the sample space with σ-field $\mathbb{F}$ and probability measure $\mathbb{P}$. Our description of the stochastic flow model follows from [96–98] where the deterministic counterpart of this problem is considered.
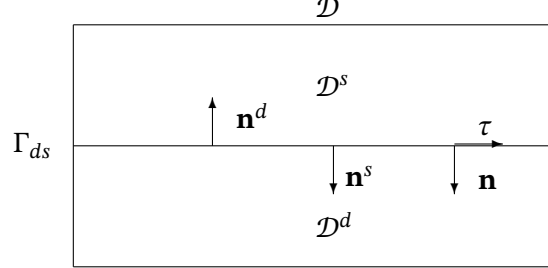

Figure 5.1: Geometry of the Darcy-Stokes problem coupled with the transport equation. Subdivision of the domain $\mathcal{D}$ into a free-flow subregion $\mathcal{D}^s$ and a porous medium subdomain $\mathcal{D}^d$, by an internal interface Γ_{ds}.

Porous medium description. The steady-state single-phase flow in a porous medium can be modeled by Darcy's law and the incompressibility condition

$$\eta k^{-1} \mathbf{u}^d + \nabla p^d = \mathbf{0} \quad \text{in} \quad \mathcal{D}^d \times \Omega, \tag{5.1}$$

$$\nabla \cdot \mathbf{u}^d = f^d \quad \text{in} \quad \mathcal{D}^d \times \Omega. \tag{5.2}$$

The fluid pressure is represented by p^d and the velocity vector $\mathbf{u}^d = (u^d, v^d)$ denotes the horizontal and vertical components. The fluid viscosity is denoted by the positive constant η and k is the spatially variable random permeability modeled as a lognormal random field, see Section 3.2. The known source (sink) term is indicated by $f^d \in L^2(\mathcal{D}^d)$, where $L^2(\mathcal{D}^d)$ is the space of square-integrable functions in $\mathcal{D}^d$.

We divide the boundary $\partial \mathcal{D}^d \backslash \Gamma_{ds}$ into two disjoint sets, $\partial \mathcal{D}^d_D$ and $\partial \mathcal{D}^d_N$, where deterministic Dirichlet and Neumann boundary conditions are prescribed,

$$p^d = g^d_D \quad \text{on} \quad \partial \mathcal{D}^d_D, \tag{5.3}$$

$$\mathbf{u}^d \cdot \mathbf{n} = g^d_N \quad \text{on} \quad \partial \mathcal{D}^d_N, \tag{5.4}$$

respectively, with $\mathbf{n}$ the outward normal to the boundary $\partial \mathcal{D}^d_N$ and $g^d_D \in L^2(\partial \mathcal{D}^d_D), g^d_N \in L^2(\partial \mathcal{D}^d_N)$.

We will operate under the assumption that the solid framework is rigid and there is no interaction between the fluid and the solid matrix of the porous medium.

Free-flow description. We consider a creeping, steady-state flow connected to the porous flow regime. We assume a viscous, incompressible Newtonian fluid flow which can be modeled by the Stokes equations, consisting of the momentum and continuity equations

$$-\nabla \cdot \sigma^s = \mathbf{f}^s \quad \text{in} \quad \mathcal{D}^s \times \Omega, \tag{5.5}$$

$$\nabla \cdot \mathbf{u}^s = \mathbf{0} \quad \text{in} \quad \mathcal{D}^s \times \Omega, \tag{5.6}$$

respectively, with $\mathbf{u}^s = (u^s, v^s)$ as the fluid velocity components and $\mathbf{f}^s = (f_1^s, f_2^s) \in L^2(\mathcal{D}^s)^2$ is the force. The fluid stress tensor $\boldsymbol{\sigma}^s$ is defined by

$$\boldsymbol{\sigma}^s \quad = \quad -p^s \mathbf{I} + 2\eta \mathbf{R}(\mathbf{u}^s), \tag{5.7}$$

$$\mathbf{R} \quad := \quad \frac{1}{2}(\nabla \mathbf{u}^s + (\nabla \mathbf{u}^s)^T), \tag{5.8}$$

where p^s denotes the fluid pressure and $\mathbf{R}$ is the strain tensor. In 2D, the components of the stress tensor (5.7) are given by

$$\boldsymbol{\sigma}^s := \begin{bmatrix} \sigma_{xx}^s & \sigma_{xy}^s \\ \sigma_{yx}^s & \sigma_{yy}^s \end{bmatrix}, \tag{5.9}$$

and they are related to the primitive variables via

$$\sigma_{xx}^s \quad = \quad -p^s + 2\eta \frac{\partial u^s}{\partial x}, \tag{5.10}$$

$$\sigma_{xy}^s \quad = \quad \sigma_{yx}^s \quad = \quad \eta \left(\frac{\partial u^s}{\partial y} + \frac{\partial v^s}{\partial x} \right), \tag{5.11}$$

$$\sigma_{yy}^s \quad = \quad -p^s + 2\eta \frac{\partial v^s}{\partial y}. \tag{5.12}$$

The exterior boundary $\partial \mathcal{D}^s \backslash \Gamma_{ds}$ is partitioned into two disjoint sets, $\partial \mathcal{D}_D^s$ and $\partial \mathcal{D}_N^s$, and the free-flow model is completed by imposing the following boundary conditions

$$\mathbf{u}^s \quad = \quad \mathbf{g}_D^s \quad \text{on} \quad \partial \mathcal{D}_D^s, \tag{5.13}$$

$$\boldsymbol{\sigma}^s \cdot \mathbf{n} \quad = \quad \mathbf{g}_N^s \quad \text{on} \quad \partial \mathcal{D}_N^s, \tag{5.14}$$

where $\mathbf{n}$ is the outward normal to the boundary $\partial \mathcal{D}_N^s$ and $\mathbf{g}_D^s \in L^2(\partial \mathcal{D}_D^s)^2$, $\mathbf{g}_N^s \in L^2(\partial \mathcal{D}_N^s)^2$.

Interface conditions. The coupling between the two problems is attained by imposing three interface conditions for mass conservation, normal stress balance and the third condition relating the slip velocity to the shear stress along Γ_{ds}

$$\mathbf{u}^s \cdot \mathbf{n}^s \quad = \quad \mathbf{u}^d \cdot \mathbf{n}^s \quad \text{on} \quad \Gamma_{ds} \times \Omega, \tag{5.15}$$

$$-\mathbf{n}^s \cdot \boldsymbol{\sigma}^s \cdot \mathbf{n}^s \quad = \quad p^d \quad \text{on} \quad \Gamma_{ds} \times \Omega, \tag{5.16}$$

$$\mathbf{u}^s \cdot \boldsymbol{\tau} + \left(\frac{\sqrt{k}}{\alpha_{BJ}} \right) \boldsymbol{\tau} \cdot \boldsymbol{\sigma}^s \cdot \mathbf{n}^s \quad = \quad \mathbf{0} \quad \text{on} \quad \Gamma_{ds} \times \Omega, \tag{5.17}$$

respectively. We denote by $\boldsymbol{\tau}$ and $\mathbf{n}^s$ the unit tangential and normal vectors to the interface Γ_{ds}, respectively. The third interface condition (5.17) was originally derived by Beaver and Joseph [85] on the basis of experimentation and dimensional analysis, and has been mathematically proven by Saffmann [86]. It is often referred to as the Beaver-Joseph-Saffmann (BJS) condition. The literature on BJS is growing rapidly as it is well able to represent the physics along the interface of viscous fluid flow and a porous media, for more details see [92, 99, 100]. Note that friction constant $\sqrt{k}/\alpha_{BJ}$ depends on the permeability along the interface and is thus here a random variable. Parameter $\alpha_{BJ} > 0$ is a dimensionless quantity measured experimentally and can also be prone to uncertainty.

Alternatively, one obtains a *no-slip interface* condition by neglecting the second term from (5.17)

$$\mathbf{u}^s \cdot \boldsymbol{\tau} = \mathbf{0} \quad \text{on} \quad \Gamma_{ds} \times \Omega. \tag{5.18}$$

Transport model. The generic single-component transport equation with random initial data gives us the following equation:

$$\phi \frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{u} - \mathbb{D}\nabla c) \quad = \quad \phi f^t \quad \text{in} \quad \mathcal{D} \times \mathcal{T} \times \Omega, \tag{5.19}$$

$$c \quad = \quad c_0 \quad \text{in} \quad \mathcal{D} \times \Omega, \quad t = 0. \tag{5.20}$$

where c denotes the concentration of the chemical component, typically expressed in terms of moles per unit volume; $\phi \in (0,1]$ is the known porosity of the medium while $f^t \in L^2(\mathcal{D})$ is a net volumetric source for c. The initial concentration $c_0 := c_0(\mathbf{x}, \omega) \in L^2(\mathcal{D})$ corresponds to a random field. The steady-state velocity field $\mathbf{u} = \mathbf{u}^s \cup \mathbf{u}^d$ is derived from the solution of the coupled Darcy-Stokes problem. In the Darcy domain, the hydrodynamical dispersion is represented by the tensor field $\mathbb{D}$ accounting for the molecular diffusion and mechanical dispersion. Molecular diffusion, which takes place due to a concentration gradient, is more significant compared to the mechanical dispersion resulting from micro-scale variations in the velocity field. In this chapter, we consider a diffusion tensor as defined in [101, 102] for an isotropic porous media with tensor components :

$$D_{xx} = D_L \frac{u^2}{|\mathbf{u}|} + D_T \frac{v^2}{|\mathbf{u}|} + D^*, \tag{5.21}$$

$$D_{yy} = D_L \frac{v^2}{|\mathbf{u}|} + D_T \frac{u^2}{|\mathbf{u}|} + D^*, \tag{5.22}$$

$$D_{xy} = D_{yx} = (D_L - D_T)\frac{uv}{|\mathbf{u}|}. \tag{5.23}$$

where $D_L, D_T > 0$ are longitudinal and transverse dispersivity, respectively, and $D^* > 0$ is the effective molecular diffusion. In the Stokes domain, the dispersion tensor is usually taken to be isotropic, i.e. $\mathbb{D} = D\mathbf{I}$.

The transport model is completed by applying a Cauchy boundary condition at the inflow boundary and a non-dispersive mass flux condition at outflow boundaries, formally expressed as

$$(c\mathbf{u} - \mathbb{D}\nabla c) \cdot \mathbf{n} = (c_{in}\mathbf{u}) \cdot \mathbf{n} \quad \partial\mathcal{D}_{in} \times \mathcal{T}, \tag{5.24}$$

$$\mathbb{D}\nabla c \cdot \mathbf{n} = 0 \quad \partial\mathcal{D}_{out} \times \mathcal{T}, \tag{5.25}$$

respectively. The inflow boundary is defined as $\partial\mathcal{D}_{in} := \{\mathbf{x} \in \partial\mathcal{D} : \mathbf{u} \cdot \mathbf{n} < 0\}$ and the outflow boundary as $\partial\mathcal{D}_{out} := \partial\mathcal{D}\backslash\partial\mathcal{D}_{in}$.

For ease of presentation, we consider a simplified chemical transport model ignoring the adsorption phenomena and also assume that the injection of the chemical components does not have any effect on the steady-state Darcy-Stokes flow field.

5.3. Finite volume discretization

For the spatial discretizations, we employ the finite volume scheme on a staggered grid for the Darcy-Stokes system. The domain is subdivided into square blocks of size $h \times h$ conforming with $\partial\mathcal{D}$ and Γ_{ds}. The locations and indexing of unknowns u, v and p, c along with their respective control volumes $\mathcal{D}_h^1, \mathcal{D}_h^2$ and $\mathcal{D}_h^3$ are shown in figures 5.2 - 5.3.

As the same variables describe different physics in the two subdomains, discretization along the interface becomes involved. A large numerical error or even a reduction in the order of grid convergence may be encountered if interface conditions are not handled properly. A staggered arrangement of the unknowns greatly simplifies the discretization along the interface and has been proven to be effective in reducing numerical error along the interface [98]. Moreover, a staggered grid is also a convenient way of avoiding spurious oscillations in the numerical solution [103] and obtaining conservation of mass throughout the system, also on a relatively coarse grid.

To make this chapter self-contained, we briefly discuss the spatio-temporal discretization of the coupled problem. For the Darcy-Stokes approximation, we will closely follow the description from [98].
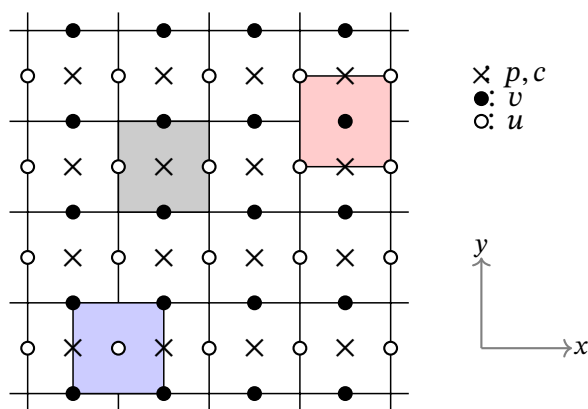
Figure 5.2: Staggered grid location of unknown and corresponding control volumes.
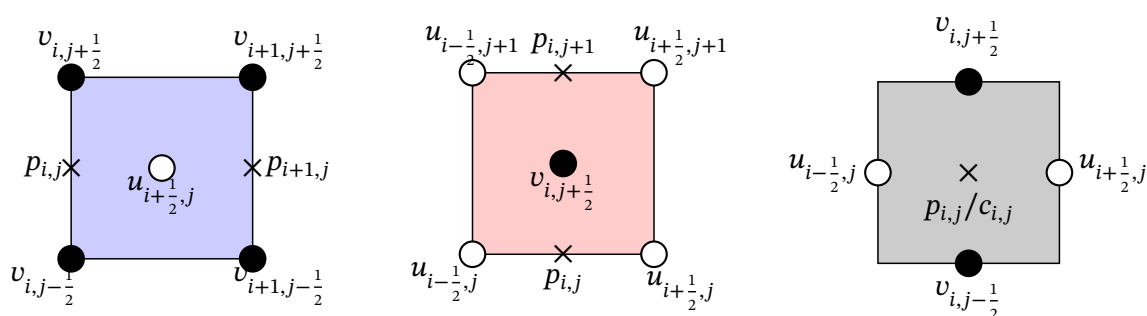

Figure 5.3: Control volumes $\mathcal{D}_h^1(i+1/2, j)$ (left), $\mathcal{D}_h^2(i, j+1/2)$ (middle) and $\mathcal{D}_h^3(i, j)$ (right) for the primary unknowns u, v and p, c respectively, together with the corresponding indexing for each variable.

5.3.1   Discretization of Darcy-Stokes flow

Discretization of Darcy equation

The discrete equations for velocities and pressure in the mixed formulation are easy to determine. The discretization for the horizontal velocity is obtained by integrating the Darcy equation (5.1) in the control volume $\mathcal{D}_h^1(i+1/2, j)$

$$\eta k_{i+\frac{1}{2}, j}^{-1} u_{i+\frac{1}{2}, j}^d + \frac{p_{i+1,j}^d - p_{i,j}^d}{h} = 0. \tag{5.26}$$

Similarly, the discrete equations for the vertical velocities are obtained by integrating the Darcy equation in the control volume $\mathcal{D}_h^2(i, j+1/2)$. For the pressure, we integrate the continuity equation (5.2) in the volume $\mathcal{D}_h^3(i, j)$

$$\frac{u_{i+\frac{1}{2}, j}^d - u_{i-\frac{1}{2}, j}^d}{h} + \frac{v_{i,j+\frac{1}{2}}^d - v_{i,j-\frac{1}{2}}^d}{h} = f_{i,j}^d. \tag{5.27}$$

For boundaries where the pressure is prescribed see Eq. (5.3), we integrate the Darcy equation over half volumes. At boundaries for which the velocities are known, (5.4) is directly applied.

Discretization of Stokes equation

Integrating the first component of the momentum equation (5.5) in the control volumes $\mathcal{D}_h^2(i, j + 1/2)$, gives

$$-\left(\frac{(\sigma_{xx})_{i+1,j} - (\sigma_{xx})_{i,j}}{h} + \frac{(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} - (\sigma_{xy})_{i+\frac{1}{2},j-\frac{1}{2}}}{h}\right) = (f_1^s)_{i+\frac{1}{2},j}. \tag{5.28}$$

The components in the above equation are approximated using (5.10) as

$$(\sigma_{xx})_{i+1,j} = -p_{i+1,j}^s + 2\eta \frac{u_{i+\frac{3}{2},j}^s - u_{i+\frac{1}{2},j}^s}{h}, \tag{5.29}$$

$$(\sigma_{xx})_{i,j} = -p_{i,j}^s + 2\eta \frac{u_{i+\frac{1}{2},j}^s - u_{i-\frac{1}{2},j}^s}{h}, \tag{5.30}$$

$$(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} = \eta\left(\frac{u_{i+\frac{1}{2},j+1}^s - u_{i+\frac{1}{2},j}^s}{h} + \frac{v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h}\right), \tag{5.31}$$

$$(\sigma_{xy})_{i+\frac{1}{2},j-\frac{1}{2}} = \eta\left(\frac{u_{i+\frac{1}{2},j}^s - u_{i+\frac{1}{2},j-1}^s}{h} + \frac{v_{i+1,j-\frac{1}{2}}^s - v_{i,j-\frac{1}{2}}^s}{h}\right). \tag{5.32}$$

Using these approximations in (5.28), we finally get

$$-\frac{2\eta}{h^2}(u_{i+\frac{3}{2},j}^s - 2u_{i+\frac{1}{2},j}^s + u_{i-\frac{1}{2},j}^s) - \frac{\eta}{h^2}(u_{i+\frac{1}{2},j+1}^s - 2u_{i+\frac{1}{2},j}^s + u_{i+\frac{1}{2},j-1}^s)$$

$$-\frac{\eta}{h^2}(v_{i+1,j+\frac{1}{2}}^s - v_{i,j+\frac{1}{2}}^s - v_{i+1,j-\frac{1}{2}}^s + v_{i,j-\frac{1}{2}}^s) + \frac{1}{h}(p_{i+1,j}^s - p_{i,j}^s) = (f_1^s)_{i+\frac{1}{2},j}. \tag{5.33}$$

Similarly, the second component of the momentum equation is integrated in the volume $\mathcal{D}_h^2(i, j + 1/2)$. The continuity equation (5.6) is as given in (5.37) over the volume $\mathcal{D}_h^3(i, j)$.

The Dirichlet boundary condition (5.13) can be directly utilized for the approximation of the stress component in (5.29). If the stress components are prescribed along the boundary (5.14) then the Stokes equations are integrated over half volumes along that boundary.

Discretization of interface equations

An appropriate interface discretization is crucial for achieving a strong numerical coupling between two sub-solutions. As the vertical velocities lie along the interface, we integrate the Stokes equation for v^s over the half volume indicated in red in Figure 5.4,

$$-\left(\frac{(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} - (\sigma_{xy})_{i-\frac{1}{2},j+\frac{1}{2}}}{h} + \frac{(\sigma_{yy})_{i,j+1} - (\sigma_{yy})_{i,j+\frac{1}{2}}}{h/2}\right) = (f_2^s)_{i,j+\frac{1}{2}}. \tag{5.34}$$

Now, we describe the computation of the stress components in the above equation. To approximate these terms, we will utilize the interface conditions (5.15)-(5.17).

For the normal stress component $(\sigma_{yy})_{i,j+1}$, we use

$$(\sigma_{yy})_{i,j+1} = -p_{i,j+1}^s + 2\eta \frac{v_{i,j+\frac{3}{2}}^s - v_{i,j+\frac{1}{2}}^s}{h}. \tag{5.35}$$

Next, the normal stress $(\sigma_{yy})_{i,j+1/2}$ at the interface is derived by equilibrating the normal stress with the pressure at the interface

$$(\sigma_{yy})_{i,j+\frac{1}{2}} = p^d_{i,j+\frac{1}{2}}. \tag{5.36}$$

Pressure $p^d_{i,j+1/2}$ at the interface is not known and is obtained by integrating the Darcy equation over the half volume indicated by the gray box in Figure 5.4, as

$$\eta k^{-1}_{i,j+\frac{1}{2}} v^d_{i,j+\frac{1}{2}} + \frac{p^d_{i,j+\frac{1}{2}} - p^d_{i,j}}{h/2} = 0. \tag{5.37}$$

Equation (5.36) can then be rewritten as

$$(\sigma_{yy})_{i,j+\frac{1}{2}} = -p^d_{i,j} + \frac{\eta h}{2k_{i,j+\frac{1}{2}}} v^d_{i,j+\frac{1}{2}}. \tag{5.38}$$

The stress component $(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}}$ is approximated as

$$(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} = \eta \left(\frac{u^s_{i+\frac{1}{2},j+1} - u^s_{i+\frac{1}{2},j+\frac{1}{2}}}{h/2} + \frac{v^s_{i+1,j+\frac{1}{2}} - v^s_{i,j-\frac{1}{2}}}{h} \right). \tag{5.39}$$

Here, the horizontal component of the velocity at the interface $u^s_{i+1/2,j+1/2}$ is derived using the BJS condition (5.17). Therefore,

$$\left(\frac{\sqrt{k_{i,j+\frac{1}{2}}}}{\alpha_{BJ}} \right)^{-1} u^s_{i+\frac{1}{2},j+\frac{1}{2}} - \eta \left(\frac{u^s_{i+\frac{1}{2},j+1} - u^s_{i+\frac{1}{2},j+\frac{1}{2}}}{h/2} + \frac{v^s_{i+1,j+\frac{1}{2}} - v^s_{i,j+\frac{1}{2}}}{h} \right) = 0. \tag{5.40}$$

Using (5.39) and (5.40), one obtains:

$$(\sigma_{xy})_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{2\eta \widetilde{m}_{i,j+\frac{1}{2}}}{h} u^s_{i+\frac{1}{2},j+1} + \eta \widetilde{m}_{i,j+\frac{1}{2}} \frac{v^s_{i+1,j+\frac{1}{2}} - v^s_{i,j+\frac{1}{2}}}{h}, \tag{5.41}$$

where

$$\widetilde{m}_{i,j+\frac{1}{2}} = \left(1 - \frac{2\eta \sqrt{k_{i,j+\frac{1}{2}}}}{h\alpha_{BJ} + 2\eta \sqrt{k_{i,j+\frac{1}{2}}}} \right).$$

The final component $(\sigma_{xy})_{i-\frac{1}{2},j+\frac{1}{2}}$ is computed in a similar manner. Using all stress components derived above, we rewrite (5.34) as

$$- \frac{2\eta \widetilde{m}_{i,j+\frac{1}{2}}}{h^2} (u^s_{i+\frac{1}{2},j+1} - u^s_{i-\frac{1}{2},j+1}) - \frac{\eta \widetilde{m}_{i,j+\frac{1}{2}}}{h^2} (v^s_{i+1,j+\frac{1}{2}} - 2v^s_{i,j+\frac{1}{2}} + v^s_{i-1,j+\frac{1}{2}})$$

$$+ \frac{2}{h} (p^s_{i,j+1} - p^d_{i,j}) - \frac{4\eta}{h^2} (v^s_{i,j+\frac{3}{2}} - v^s_{i,j+\frac{1}{2}}) + \frac{v^s_{i,j+\frac{1}{2}}}{k_{i,j+\frac{1}{2}}} = (f^s_2)_{i,j+\frac{1}{2}}. \tag{5.42}$$
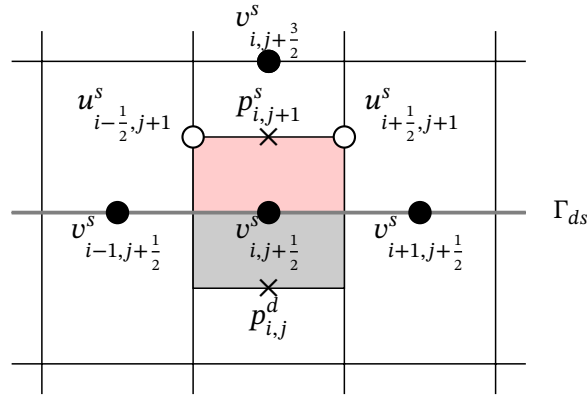
Figure 5.4: Locations of unknowns required for discretization along the interface.

5.3.2 Discretization of transport equation

The numerical scheme to approximate the advection-diffusion part of the transport equation depends on the type of problem. In the case of contaminant transport in the Darcy-Stokes system, convection is the dominant cause for the flow movement [101, 102]. An important requirement for designing an efficient MLMC estimator is using numerical schemes that are stable on coarse meshes. For example, for time-stepping an implicit method is favourable due to unconditional stability with respect to time-step size. Similarly, an upwind based discretization can give stable solutions on relatively coarse grids compared to the central differencing based methods. We briefly discuss the implementation details of the spatio-temporal discretization of the transport equation.

In each of the control volumes $\mathcal{D}_h^3(i,j)$ (see Figure 5.3), the integral formulation of (5.19) takes the form

$$\frac{\partial}{\partial t} \int_{\mathcal{D}_h^3(i,j)} \phi c \mathrm{d}\mathbf{x} + \int_{\mathcal{D}_h^3(i,j)} \nabla \cdot (c\mathbf{u} - \mathbb{D}\nabla c)\mathrm{d}\mathbf{x} = \int_{\mathcal{D}_h^3(i,j)} \phi f^t \mathrm{d}\mathbf{x}. \tag{5.43}$$

The velocity field $\mathbf{u}$ is assumed to be exact but for the current problem this is derived from the solution of coupled Darcy-Stokes system. Also, for simplicity, we will assume $D_{xy} = D_{yx} = 0$ resulting in a diagonal dispersion tensor $\mathbb{D}$.

5.3.3 Spatial discretization

Using the Gauss divergence theorem, the second integral in (5.43) is reformulated as a boundary integral for the boundary $\partial \mathcal{D}_h^3(i,j) = \bigcup_{\kappa=1}^4 \partial \mathcal{D}_h^3(i,j,\kappa)$, such that

$$\int_{\mathcal{D}_h^3(i,j)} \nabla \cdot (c\mathbf{u} - \mathbb{D}\nabla c)\mathrm{d}\mathbf{x} = \sum_{\kappa=1}^4 \int_{\partial \mathcal{D}_h^3(i,j,\kappa)} (c\mathbf{u} - \mathbb{D}\nabla c) \cdot \mathbf{n}_\kappa \mathrm{d}S_\kappa, \tag{5.44}$$

where $\mathbf{n}_\kappa$ denotes the unit normal vector to the corresponding face. Flow in the Stokes domain is well-behaved and the central differencing scheme to evaluate the fluxes in (5.44) will result in a second-order accurate solution locally on very fine grids. However, in the Darcy domain, a permeability field with very small correlation lengths may exhibit a highly fluctuating flow and may give rise to non-physical oscillations in the solution even on finer meshes. The quality of solution

will depend on the *cell Péclet number* which is defined as

$$Pe_h(i,j) := \min \left\{ \left| \frac{u}{D_{xx}} \right|_{i-\frac{1}{2},j}, \left| \frac{u}{D_{xx}} \right|_{i+\frac{1}{2},j}, \left| \frac{v}{D_{yy}} \right|_{i,j-\frac{1}{2}}, \left| \frac{v}{D_{yy}} \right|_{i,j+\frac{1}{2}} \right\} h. \tag{5.45}$$

To avoid oscillations it is desired to have $Pe_h(i,j) \leq 2$ in each cell. For the current problem, the velocities u_h, v_h are random variables and it is difficult to bound the cell Péclet number, especially on coarser grids. A first-order Upwind Differencing Scheme (UDS) for the approximation of the convective fluxes seems a convenient choice. The main disadvantage of the first-order upwind scheme is however excessive numerical diffusion that smears out sharp features in the solution. Therefore, to prevent this we use a Quadratic Upwind Interpolation for Convective Kinematics (QUICK) [104] scheme for the approximation of the convective fluxes. For the diffusion part we will use the central differencing scheme. This scheme is less diffusive compared to the first-order solves and is also highly stable. Furthermore, the QUICK scheme is formally third-order accurate in space thus can capture the additional smoothness in the Stokes region. The QUICK scheme is however, not monotone, which means that flux limiters will be employed for problems with sharp gradients. Many versions of the flux-limited QUICK method exist in the literature [105, 106] mostly based on the deferred correction (or sometimes referred to as defect correction) framework where the first-order scheme is improved using a higher-order correction. In this work, we consider the implicit version for the same. The flux for the face "BC" (see Figure 5.5) is defined as follows

$$\langle F_{i+\frac{1}{2},j} \rangle := \langle F_{i+\frac{1}{2},j} \rangle^{UDS} + \Psi(r_{i+\frac{1}{2},j}) \left(\langle F_{i+\frac{1}{2},j} \rangle^{QUICK} - \langle F_{i+\frac{1}{2},j} \rangle^{UDS} \right), \tag{5.46}$$

where $\langle F_{i+1/2,j} \rangle^{UDS}$ and $\langle F_{i+1/2,j} \rangle^{QUICK}$ are face-averaged convective fluxes computed using the UDS and QUICK scheme, respectively, and are computed as

$$\langle F_{i+\frac{1}{2},j} \rangle^{UDS} = \begin{cases} u_{i+\frac{1}{2},j}(c_{i,j}), & \text{if } u_{i+\frac{1}{2},j} \geq 0, \\ u_{i+\frac{1}{2},j}(c_{i+1,j}), & \text{if } u_{i+\frac{1}{2},j} < 0, \end{cases} \tag{5.47}$$

and

$$\langle F_{i+\frac{1}{2},j} \rangle^{QUICK} = \begin{cases} u_{i+\frac{1}{2},j} \left(\frac{3}{8}c_{i+1,j} + \frac{3}{4}c_{i,j} - \frac{1}{8}c_{i-1,j} \right), & \text{if } u_{i+\frac{1}{2},j} \geq 0, \\ u_{i+\frac{1}{2},j} \left(\frac{3}{8}c_{i,j} + \frac{3}{4}c_{i+1,j} - \frac{1}{8}c_{i+2,j} \right), & \text{if } u_{i+\frac{1}{2},j} < 0. \end{cases} \tag{5.48}$$

The limiter function is $\Psi(r_{i+1/2,j})$, where $r_{i+1/2,j}$ is the measure of the local smoothness. This is computed by taking the ratio of successive gradients along the stream-wise direction to decide the weights for the first-order and QUICK schemes, where

$$\tilde{r}_{i+\frac{1}{2},j} = \begin{cases} \dfrac{c_{i+1,j} - c_{i,j}}{c_{i,j} - c_{i-1,j}}, & \text{if } u_{i+\frac{1}{2},j} \geq 0, \\ \dfrac{c_{i+2,j} - c_{i+1,j}}{c_{i+1,j} - c_{i,j}}, & \text{if } u_{i+\frac{1}{2},j} < 0. \end{cases} \tag{5.49}$$

and

$$r_{i+\frac{1}{2},j} := \min \left(\tilde{r}_{i+\frac{1}{2},j}, \tilde{r}^{-1}_{i+\frac{1}{2},j} \right).$$
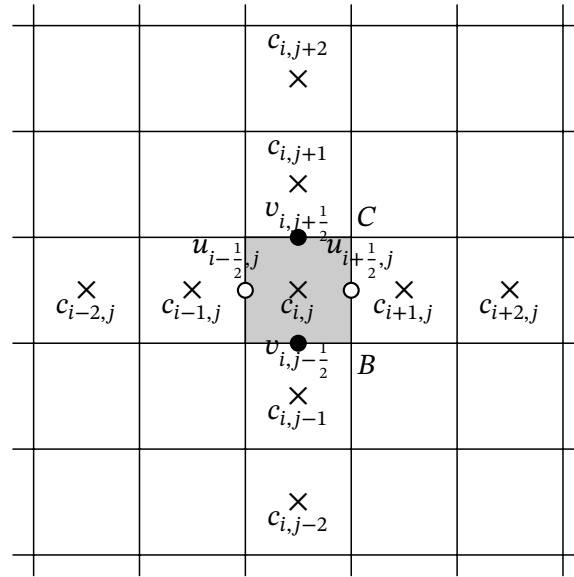
Figure 5.5: Stencil for the QUICK scheme for purely convection problem.

The inverse of $\tilde{r}$ is used to find large denominators in (5.49). The choice of the limiter function Ψ is somewhat arbitrary. We have numerically tested a number of classical TVD-limiter functions (for e.g., see [107]) and they performed reasonably well. In the current application, where we will encounter discontinuous boundary and initial conditions, the Koren limiter [108] with $\Psi(r) := \max[0, \min(2r, (2 + r)/3, 2)]$ was able to reconstruct the propagating sharp fronts. The gradients can be computed using the solution from the previous time-step and then we modify the fluxes using (5.46) to obtain the improved solution. Note that the staggered arrangement of variables for the Darcy-Stokes flow is very well suited here as the velocities at the cell faces are directly obtained from the solution of discrete Darcy-Stokes flow. The inflow and outflow boundary conditions (5.24)-(5.25) are utilized to compute fluxes for control volumes that lie along the domain boundary.

5.3.4 Temporal discretization

For the temporal discretizations, many well-established methods are available in the literature. In our application, flexibility in the time-step size is needed and therefore we prefer an implicit time-stepping scheme. We can formally rewrite the transport equation (5.19) in a semi-discrete form as

$$\frac{\partial c_h}{\partial t} + \mathcal{L}_h c_h = f_h \quad \text{in} \quad \mathcal{D}_h \times \mathcal{T} \times \Omega, \tag{5.50}$$

with $\mathcal{L}_h$ denoting the spatial discretization operator. For ease of presentation, we have omitted the porosity $\phi(\mathbf{x})$ term in (5.50) by scaling $\mathcal{L}_h$ by $\phi_{i,j}$. With the time-step size $\Delta t > 0$ and for any given integer $m \geq 1$, we define the temporal grid points as $t^m := m\Delta t$ and will denote by c_h^m the approximation of $c(\cdot, t^m)$. To solve (5.50), we propose to use an Alternating Direction Implicit based solution method that decomposes the semi-discrete form (5.50) into two linear systems each requiring inversion of a tridiagonal matrix for UDS or a pentadiagonal matrix for QUICK scheme. The ADI solver uses two half time-steps. First, an intermediate quantity $c^{m-1/2}$ is generated by performing an implicit Euler time-stepping along the x-coordinate and an explicit Euler along y-direction. Using this intermediate quantity, the reverse is performed for the next half step. Motivated

by the classical Peaceman-Rachford approach [109], one can express the ADI method as

$$\left(I_h + \tfrac{\Delta t}{2}\mathcal{L}_h^x\right)c_h^{m-1/2} = \tfrac{\Delta t}{2}f_h^{m-1/2} + \left(I_h - \tfrac{\Delta t}{2}\mathcal{L}_h^y\right)c_h^{m-1}, \tag{5.51}$$

$$\left(I_h + \tfrac{\Delta t}{2}\mathcal{L}_h^y\right)c_h^m = \tfrac{\Delta t}{2}f_h^m + \left(I_h - \tfrac{\Delta t}{2}\mathcal{L}_h^x\right)c_h^{m-1/2}, \tag{5.52}$$

where I_h is an identity matrix, $\mathcal{L}_h^y$ and $\mathcal{L}_h^x$ are tridiagonal (pentadiagonal) matrices derived from discretizations along x- and y-coordinate, respectively. There are many standard algorithms available for inverting tridiagonal or pentadiagonal matrices, for example, the Thomas algorithm can efficiently invert a tridiagonal matrix with only 8 floating point operations per unknown. Furthermore, this ADI scheme is second-order accurate in time which can be easily verified using Taylor's expansion.

5.4. Multigrid for Darcy-Stokes problem with lognormal diffusion

Multigrid methods are generally recognized as fast efficient solution methods for a large class of linear and nonlinear problems. We propose a *monolithic multigrid algorithm* to solve the coupled Darcy-Stokes system with a heterogeneous stochastic permeability field. This multigrid algorithm stems from the work done in [98], where a fixed permeability value was used throughout the Darcy domain. We extend the method to the cases with highly heterogeneous permeability fields. Solving the Darcy flow with highly fluctuating and discontinuous diffusion parameters is quite challenging. In the context of geometric multigrid methods, basically two approaches exist, based on either the cell-centered or the vertex-centered location of unknowns, see [14, 23, 54] and references therein. The vertex-centered approach requires transfer operators that are dependent on the diffusion parameter. Such multigrid implementations may be expensive as these transfer operators need to be modified on all grid levels. Due to this, the cell-centered version is somewhat more beneficial as it is possible to achieve a decent multigrid convergence using constant transfer operators [23, 54]. We will demonstrate that the cell-centered approach can be extended to the coupled Darcy-Stokes system.

We would also like to point out that another common alternative for solving the linear systems derived from a coupled model is the Domain Decomposition Method (DDM). In a DDM, one splits the main boundary value problem into smaller boundary value subproblems. Typically, the boundary information between subproblems is exchanged during every iteration until the converged solution is reached. Contrary to this, monolithic solution approaches treat the coupled system as a single problem. An important aspect of this approach is that the coupling variables between the subproblems are treated simultaneously, thus after every iteration, the three fields $(u, v, p)^T$ are updated throughout the domain. The monolithic approach has proven to be very efficient when the subproblems are strongly coupled [110, 111].

5.4.1 Uzawa smoother for saddle-point system

Discretization of each subproblem, the mixed form of Darcy flow and the Stokes equations, yields a saddle-point system. This saddle-point structure can be maintained for the coupled Darcy-Stokes system by ordering the velocity unknowns together for both subproblems followed by the pressure unknowns. This results in the following linear system

$$\begin{bmatrix} \mathcal{A}_h & \mathcal{B}_h^T \\ \mathcal{B}_h & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_h \\ p_h \end{bmatrix} = \begin{bmatrix} \mathbf{g}_h \\ f_h \end{bmatrix}, \tag{5.53}$$

where $\mathbf{u}_h = (\mathbf{u}_h^d, \mathbf{u}_h^s)^T$, $p_h = (p_h^d, p_h^s)^T$, $\mathbf{g}_h = (\mathbf{0}, \mathbf{f}_h^s)^T$ and $f_h = (f^d, 0)^T$. For both subproblems, $\mathcal{B}_h^T$ and $\mathcal{B}_h$ represent the discrete gradient and minus discrete divergence operators, respectively, and

$\mathcal{A}_h$ is the discrete representation of either the Laplacian operator $-\eta\Delta$ for the Stokes equations, or $\eta k_h^{-1} I_h$ for the Darcy equation.

The multigrid method proposed here uses a special class of relaxation methods called the Uzawa smoothers [95]. The Uzawa smoother is basically an equation-wise, decoupled smoother where the velocity components in the Darcy and Stokes domains are first updated, after which the pressure field is updated. In the following, we provide the details of this smoother.

Consider a splitting of the matrix coefficients of the saddle point system (5.53) as

$$
\begin{bmatrix} \mathcal{A}_h & \mathcal{B}_h^T \\ \mathcal{B}_h & 0 \end{bmatrix} = \begin{bmatrix} \mathcal{M}_h & 0 \\ \mathcal{B}_h & -\zeta_h^{-1} I_h \end{bmatrix} - \begin{bmatrix} \mathcal{M}_h - \mathcal{A}_h & -\mathcal{B}_h^T \\ 0 & -\zeta_h^{-1} I_h \end{bmatrix},
\tag{5.54}
$$

where $\mathcal{M}_h$ is a smoother for the operator $\mathcal{A}_h$ and ζ_h is a positive parameter. For a given approximation of the solution $(\mathbf{u}_h^{k-1}, p_h^{k-1})^T$, the relaxed approximation $(\mathbf{u}_h^k, p_h^k)^T$, can be defined in the following way:

$$
\begin{bmatrix} \mathcal{M}_h & 0 \\ \mathcal{B}_h & -\zeta_h^{-1} I_h \end{bmatrix} \begin{bmatrix} \mathbf{u}_h^k \\ p_h^k \end{bmatrix} = \begin{bmatrix} \mathcal{M}_h - \mathcal{A}_h & -\mathcal{B}_h^T \\ 0 & -\zeta_h^{-1} I_h \end{bmatrix} \begin{bmatrix} \mathbf{u}_h^{k-1} \\ p_h^{k-1} \end{bmatrix} + \begin{bmatrix} \mathbf{g}_h \\ f_h \end{bmatrix}.
\tag{5.55}
$$

In terms of velocity and pressure variables, the relaxation step can be written down as:

$$
\mathbf{u}_h^k = \mathbf{u}_h^{k-1} + \mathcal{M}_h^{-1}\left(\mathbf{g}_h - \mathcal{A}_h \mathbf{u}_h^{k-1} - \mathcal{B}_h^T p_h^{k-1}\right),
\tag{5.56}
$$

$$
p_h^k = p_h^{k-1} + \zeta_h(\mathcal{B}_h \mathbf{u}_h^k - f_h),
\tag{5.57}
$$

respectively. In [97, 98], the choice of $\mathcal{M}_h$ was based on a symmetric Gauss-Seidel iteration for (5.56) that gives

$$
\mathcal{M}_h = (\Lambda_h + L_h)\Lambda_h^{-1}(\Lambda_h + U_h),
\tag{5.58}
$$

where Λ_h, L_h and U_h are the diagonal, strictly lower and strictly upper parts of $\mathcal{A}_h$, respectively. A symmetric Gauss-Seidel iteration comprises a forward and a backward sweep for velocity in the entire domain.

Next, for smoothing of the pressure variable, a Richardson iteration (5.57) with appropriate relaxation parameters is applied. For any control volume $\mathcal{D}_h^3(i, j)$, we use the following relaxation parameters

$$
\zeta_h(i, j) = \begin{cases} \eta & \text{in} \quad \mathcal{D}^s, \\ \dfrac{\eta h^2}{5\overline{k}_h(i, j)} & \text{in} \quad \mathcal{D}^d, \end{cases}
\tag{5.59}
$$

where η is the fluid viscosity and $\overline{k}_h(i, j)$ is derived by local averaging using a *half-weighting* (HW) operator

$$
\overline{k}_h(i, j) = \frac{1}{8}[4k_{i,j} + k_{i-1,j} + k_{i+1,j} + k_{i,j-1} + k_{i,j+1}]
\tag{5.60}
$$

This is a generalization of the optimal relaxation parameters that were derived using a Local Fourier Analysis (LFA) in [97, 98]. Usually, an LFA is performed on the multigrid components obtained by freezing the coefficient field. In the case of a variable coefficient field, ζ_h needs to be modified locally. In our experience, using the weighted average (5.60) resulted in a robust convergence rate that we will later demonstrate numerically.

Remark 5.4.1 *We would like to mention that obtaining an analytic bound on the smoothing factor of the Uzawa smoother for the stochastic Darcy-Stokes flow is rather involved.*

5.4.2 Multigrid algorithm with grid partitioning

To cover realistic problems, we propose a multi-block multigrid method which is based on the grid partitioning technique [14]. A multi-block algorithm requires communication between the Stokes and Darcy domains during the multigrid iteration. In our case, the stencils do not use variables located more than one cell away from the current control volume, therefore padding the Stokes block with one extra row to store the variables from the Darcy block is sufficient to achieve the data exchange.

The multigrid hierarchy is based on uniform coarsening, i.e. the cell-width is doubled in each coarsening step. A two-block two-grid method with variable permeability field can be described by the following steps:

1. Fine grid pre-smoothing. Relax the velocities in two blocks using the symmetric Gauss-Seidel smoother. The vertical velocity from the Stokes block along the interface is transferred to the Darcy block after which the pressure is updated by performing Richardson iterations with optimal relaxation parameters given by (5.59). The updated Darcy pressure unknowns are then transferred to the Stokes overlap control volumes.

2. Defect computation and restriction. The defect (residual) is computed for each variable. Next, the residuals from the Darcy block are transferred to the Stokes overlap region. We use fixed restriction operators to transfer residuals to the coarse grid. For the velocities, we use a six-point stencil whereas for pressures a four-point restriction is applied. These stencils are denoted by

$$[R_h^{2h}]^u = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ & \star & \\ 1 & 2 & 1 \end{bmatrix}_h^{2h}, \ [R_h^{2h}]^v = \frac{1}{8} \begin{bmatrix} 1 & & 1 \\ 2 & \star & 2 \\ 1 & & 1 \end{bmatrix}_h^{2h}, \ [R_h^{2h}]^p = \frac{1}{4} \begin{bmatrix} 1 & & 1 \\ & \star & \\ 1 & & 1 \end{bmatrix}_h^{2h}, \quad (5.61)$$

respectively.

3. Coarse grid relaxation. The coarse grid defect equation is solved exactly by a direct method, e.g., Gaussian elimination. Alternatively, one can perform multiple smoothing iterations as described in step 1 to get an approximate solution on the coarse grid. We use direct discretization to obtain the discrete system on the coarse grids and the coarse grid representation of permeability k_{2h} is also obtained by using the restriction operator $[R_h^{2h}]^p$ given in (5.61). Finally, vertical Stokes velocity unknowns at the interface are transferred to the Darcy block.

4. Coarse grid correction. The error is interpolated to the fine grid, where it is used to correct the fine grid solution. The prolongation operators are chosen as the adjoint of the restriction operators.

5. Fine grid post-smoothing. The interpolation process introduces new errors in the solution that need to be eliminated. For this, several post-smoothing steps are performed with the same procedure as used in pre-smoothing given in step 1.

Although the above multi-block algorithm resembles the domain decomposition method, the main difference is that the communication between the two domains is performed on *all multigrid levels* and not just on the finest and coarsest grids. This makes the multi-block method typically more efficient than the traditional DDM.

The two-block method can be easily extended to more than two blocks that may have more interfaces without any deterioration of the multigrid convergence rate. Furthermore, this two-grid routine is easily extended to V- and W-cycling strategies. Numerical tests showed that for this problem, the W(2,2)-cycle exhibits a fast convergence rate regardless of the roughness of the coefficient fields.

5.5. Numerical experiments

We present numerical experiments for the multigrid solver and the MLMC method separately. First, we study the accuracy of the FV scheme and the convergence and robustness of the solver

with respect to the Matérn parameters using two test problems: 1) A 2-block case with the no-slip interface condition and simple boundary conditions for the Stokes and the Darcy subproblems, and 2) A realistic 4-block problem with the BJS interface condition and discontinuous boundary conditions. Here, we will also consider a non-isotropic covariance model to generate randomly layered permeability fields. Secondly, we will compute the statistical solution of the transport in Darcy-Stokes flow using the MLMC method. Furthermore, we compare the cost and accuracy of the standard MC method with the MLMC method.

5.5.1 Multigrid convergence

No-slip interface condition on 2-blocks

For the first test case, the computational domain is taken to be $\mathcal{D} \cup \partial \mathcal{D} = [0,1] \times [0,2]$, where $\mathcal{D}^d = (0,1) \times (0,1)$ and $\mathcal{D}^s = (0,1) \times (1,2)$ and the interface $\Gamma_{ds} = [0,1] \times \{1\}$. For simplicity, we consider the source terms in both domains to be zero. Figure 5.6 explains the geometry of this problem along with the boundary conditions. The inflow in the Stokes domain is described by a parabolic function $g_D^s = (y-1)(2-y)$ and tangential flow at the top and the right boundary is applied. In the Darcy region, we set tangential flow at the left and right boundaries and set $p = 0$ at the bottom boundary to mimic the effect of a gravity force. With this configuration of boundary conditions, we will have outflow only through the bottom Darcy boundary, i.e., $\Gamma_D^d \in [0,1] \times \{0\}$.
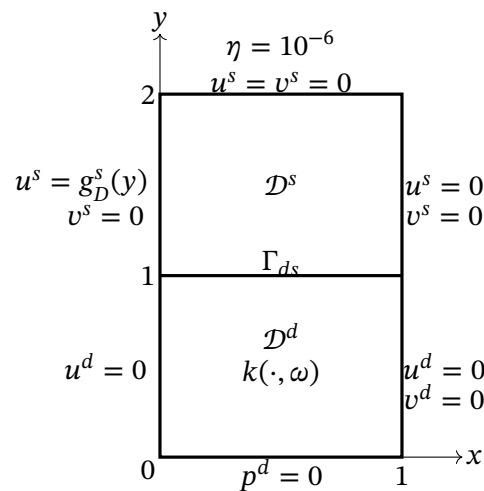

Figure 5.6: Geometry of the 2-block problem with no-slip interface condition. Subdivision of the domain into a free-flow domain $\mathcal{D}^s$ and a porous flow domain $\mathcal{D}^d$, by the interface Γ_{ds}.

The lognormal random field k is generated on a cell-centered grid and the permeability values at the face-centers are obtained using the harmonic mean of the two adjacent cell-centers, see [24] for details.

In Table 5.1, four different Matérn parameters of increasing order of complexity in terms of solvability of the linear system are presented, see also Chapter 3. Examples of random fields generated with these Matérn parameter sets were already presented in Figure 3.6. Note that for the "hardest" parameter set Φ_4, jumps up to 10 order of magnitude can be encountered.

As the performance of multigrid will depend on the particular realization of the random permeability field, we take a statistical approach to study the convergence and robustness of the multigrid algorithm. For a fixed Matérn parameter set Φ and h, we generate N_{MG} samples of the random field and record the multigrid convergence factors $\{\rho_i\}_{i=1}^{N_{MG}}$ for each solve. The samples of convergence

Table 5.1: Different combinations of the Matérn reference parameters $\Phi = (\nu_c, \lambda_c, \sigma_c^2)$ with increasing complexity from left to right.

Φ_1	Φ_2	Φ_3	Φ_4
(1.5,0.3,1)	(0.5,0.3,1)	(1.5,0.1,3)	(0.5,0.1,3)

factor ρ_i for the i-th realization of the random field are defined as

$$\rho_i := \left\{ \frac{||res^{\kappa_i}||_\infty}{||res^0||_\infty} \right\}^{\frac{1}{\kappa_i}}, \qquad \text{for} \qquad i = 1, 2, ..., N_{MG}, \tag{5.62}$$

where $||res^0||_\infty$ is the infinity-norm of the residual from an initial solution and $||res^{\kappa_i}||_\infty$ is the residual after κ_i iterations of the multigrid cycle. We use these to compute the mean multigrid convergence as

$$\langle \rho \rangle = \frac{1}{N_{MG}} \sum_{i=1}^{N_{MG}} \rho_i. \tag{5.63}$$

We compute the mean and variance of the convergence factors for the V- and W-cycles with different combinations of the number of pre- and post-smoothing steps. Similarly, the average number of multigrid iterations $\langle \kappa \rangle$ is computed. We choose random numbers as the initial solution, and the stopping criterion is reached when the κ_i-th iteration satisfies

$$\frac{||res^{\kappa_i}||_\infty}{||res^0||_\infty} \leq \varepsilon_{MG}, \tag{5.64}$$

with the tolerance ε_{MG} set to 10^{-10}. The relaxation parameters for the Richardson iteration ζ are taken as described in (5.59). In the Stokes domain $\zeta_h = \eta = 10^{-6}$ and in the Darcy domain, we take spatially varying $\zeta_h(i, j) = \eta h^2 / 5\bar{k}_h(i, j)$, where i, j are cell indices.

In tables 5.2 - 5.5, we list the average number of multigrid iterations $\langle \kappa \rangle$ (rounded up to the nearest integer) for different combinations of W-cycles required to achieve the stopping criterion on different grids for $N_{MG} = 100$. The V-cycle variants performed poorly for all Matérn parameters and also divergence was observed in case of Φ_4, hence they are not included in the tables. Also, we omit results with other values of the fluid viscosity as it did not influence the convergence factors of these cycles.

As expected, we see an increase in the number of iterations with respect to the complexity of the random field. For all Φ, we see an improvement in the convergence rate with respect to the size of the grid, stabilizing after a certain h. The stabilization is encountered on a relatively coarse grid ($1/h = 32$) for Φ_1 and on a relatively fine grid ($1/h = 128$) for Φ_4. This happens because finer grids are better able to resolve the small-scale correlation structures in the permeability field.

We use the four-point restriction operator $[R_h^{2h}]^p$ from (5.61) to get the coarse grid representation of the permeability field by which we derive the discretization operator. For a random field with small correlation length λ_c and large variance σ_c^2 that produces highly fluctuating fields, the coarse grid representation may be very different from the finest grid version. This discrepancy further increases as we go to more coarser levels. Therefore, it makes more sense to use a W-cycle which by definition performs more smoothing steps on coarse and intermediate grids where multigrid convergence may not be optimal.

Table 5.2: Average multigrid iterations $\langle \kappa \rangle$ for Φ_1

h	W(1,1)	W(1,2)	W(2,2)
1/8	36	21	16
1/16	34	20	15
1/32	35	20	15
1/64	34	19	14
1/128	33	19	14
1/256	33	19	14

Table 5.3: Average multigrid iterations $\langle \kappa \rangle$ for Φ_2

h	W(1,1)	W(1,2)	W(2,2)
1/8	35	23	17
1/16	34	21	16
1/32	35	21	16
1/64	35	21	15
1/128	35	20	15
1/256	35	20	15

Table 5.4: Average multigrid iterations $\langle \kappa \rangle$ for Φ_3

h	W(1,1)	W(1,2)	W(2,2)
1/8	60	41	32
1/16	66	43	34
1/32	53	34	27
1/64	42	25	20
1/128	40	22	16
1/256	37	21	16

Table 5.5: Average multigrid iterations $\langle \kappa \rangle$ for Φ_4

h	W(1,1)	W(1,2)	W(2,2)
1/8	59	41	33
1/16	66	42	32
1/32	57	37	28
1/64	46	30	23
1/128	40	25	18
1/256	39	23	17

It is worthwhile to mention that for Φ_1 and Φ_2, we are able to achieve the same efficiency as was observed for constant permeability in [98], which required around 14 iterations to reduce residual by 10 orders of magnitude independent of the mesh size. This is typically regarded as h-independent convergence of a multigrid algorithm. For Φ_3 and Φ_4, this efficiency is achieved on relatively finer grids.
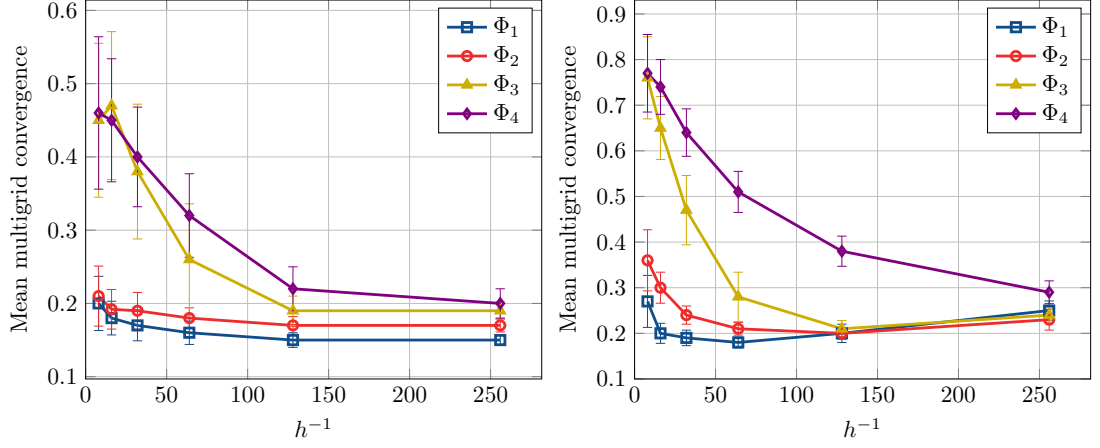
Figure 5.7: Mean convergence rate $\langle q \rangle$ along with the standard deviation for W(2,2)-cycle when k is sampled on a cell-centered grid (left) and staggered grid (right).

We wish to emphasize that the choice of sampling mesh for permeability field has a significant influence on the multigrid convergence. To study this, we also generate a permeability field on a staggered mesh such that the cell faces have different values from the cell-centers. This introduces more variability within a control volume and slows down the multigrid convergence. The mean convergence rate $\langle \rho \rangle$ along with the standard deviation for W(2,2)-cycles is presented in Figure 5.7 for cell-centered (left) and staggered (right) grid-based sampling, respectively. Multigrid convergence for a cell-centered based sampling performs well on very coarse grids and the average convergence factors on finer grids are in the range $0.1 - 0.2$. Whereas for the staggered grid-based sampling, we see a deterioration in the multigrid convergence, especially on coarser grids for parameters Φ_3 and Φ_4. As the MLMC estimator requires a relatively large number of samples on coarser grids, a cell-centered based sampling grid for k may give rise to a significant gain over the staggered sampling approach. We also remark that the FV error convergence rate does not deteriorate when a cell-centered approach is adopted which will be demonstrated later.

The most efficient variant of the W-cycle is determined on the basis of the average CPU times required to reach the stopping criterion. In Figure 5.8 we plot the CPU times for different W-cycles versus the grid size for the "hardest" parameter set Φ_4. The W(2,2)-cycle was the fastest to achieve the stopping criterion. A similar trend is also observed for the other Matérn parameters. We would like to point out that the average CPU time for one W(2,2)-cycle is about 0.65 seconds for a 256×512 grid. For many engineering applications, however, residual reduction of about 3 orders of magnitude may be sufficient and can be achieved in less than 8 iterations using the W(2,2)-cycle. Also, we observe the $\mathcal{O}(h^{-2})$ scaling of the computational cost from $h = 1/128$ onwards for all the three variants of W-cycles. As the cost is proportional to the number of unknowns, we can conclude that we have an optimal multigrid solver for the coupled Darcy-Stokes problem, on staggered meshes. We will use the W(2,2)-cycle to solve the Darcy-Stokes system for the MC samples. Next, we numerically demonstrate the FV error convergence in the mean for the coupled Darcy-Stokes approximation. In Figures 5.9, the relative errors in the horizontal and vertical components of the velocity for $\mathcal{D}^s$, $\mathcal{D}^d$ and $\mathcal{D}$, respectively, are listed. These relative errors are computed using the Monte Carlo method with a sufficient number of samples such that the sampling error is less than the FV bias. In the Stokes region, a convergence close to second-order is observed for the horizontal velocity component, however, for the vertical velocity component, we see a dependence on the smoothness parameter of the Matérn function ν_c. This is due to the fact that for the no-slip boundary condition $u^s = 0$ is imposed along the interface without any influence of the permeability field. On the other
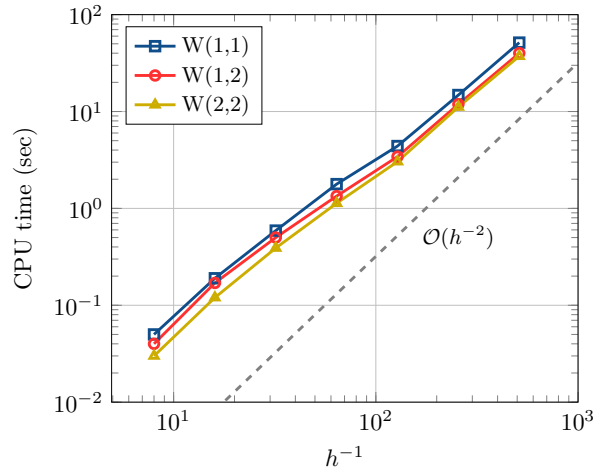
Figure 5.8: Comparison of average CPU time for W-cycle variants for Φ_4. A similar trend is observed for other Matérn parameters.

hand, the vertical velocity component at the interface is based on the discretization of the Darcy flow (5.26), thus the error in v^s depends on the smoothness of the random permeability field. For the BJS interface condition, the convergence of u^s will again depend on the smoothness of permeability field, due to (5.40).

For the Darcy domain, the FV error decays with a rate equal to ν_c for both velocity components. The global error also shows a similar trend as the dominating error is due to the approximation of the Darcy subproblem. In general, the FV error increases with an increase in the variance parameter σ_c^2.

Beaver-Joseph-Saffman interface condition on 4-blocks

Next, we test our method for a 4-block geometry with the BJS boundary condition at the Darcy-Stokes interface. This type of geometry is sometimes utilized to model a cross-flow filtration process [112]. The geometry and boundary constraints are depicted in Figure 5.10. Here, $\mathcal{D}^s$ represents a channel divided into 3-blocks and $\mathcal{D}^d$ represents a porous media forming an interface with the second Stokes block $\mathcal{D}_2^s$. We impose a constant inflow velocity $u^s = 0.1$ at the left inlet boundary of the channel. On the right, there is an exit of length $1/6$ with zero normal stress $\sigma_{xx} = \sigma_{xy} = 0$.

We also analyze the performance of the multigrid algorithm with grid partitioning (see Section 5.4.2) when extended to more than two blocks. For the 4-block test case, the algorithm can be described in the following way. Given an initial approximation, we first update all variables on the first subgrid $\mathcal{D}_1^s$. Next, the updated boundary data from $\mathcal{D}_1^s$ is communicated to the second block $\mathcal{D}_2^s$ and then all the interior points are updated. Similarly, the variables in blocks $\mathcal{D}_3^s$ and $\mathcal{D}^d$ are updated. For each block, it is common to extend the computational mesh by one cell length along the boundary between the two neighbouring blocks in order to store the updated data. Note that this grid partitioning based multigrid is also highly parallelizable as all the blocks can be simultaneously relaxed on separate processors.

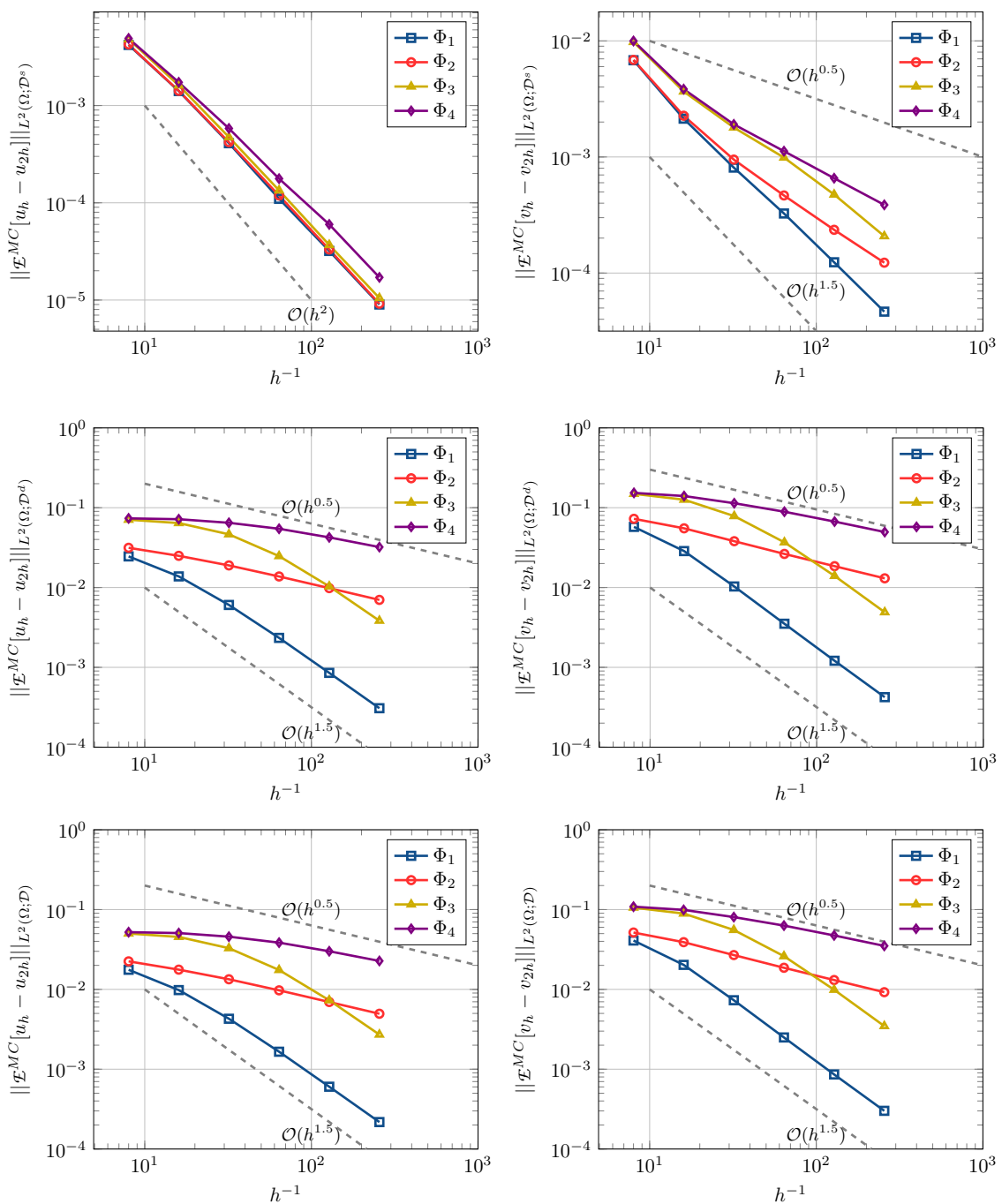To simulate layered porous media, we use a simplified version of the Matérn model described

Figure 5.9: Mean error convergence of horizontal (left column) and vertical (right column) velocity components for parameter sets $\Phi_1 - \Phi_4$ in Stokes domain $\mathcal{D}^s$ (top row), Darcy domain $\mathcal{D}^d$ (middle row) and whole domain $\mathcal{D}$ (bottom row).
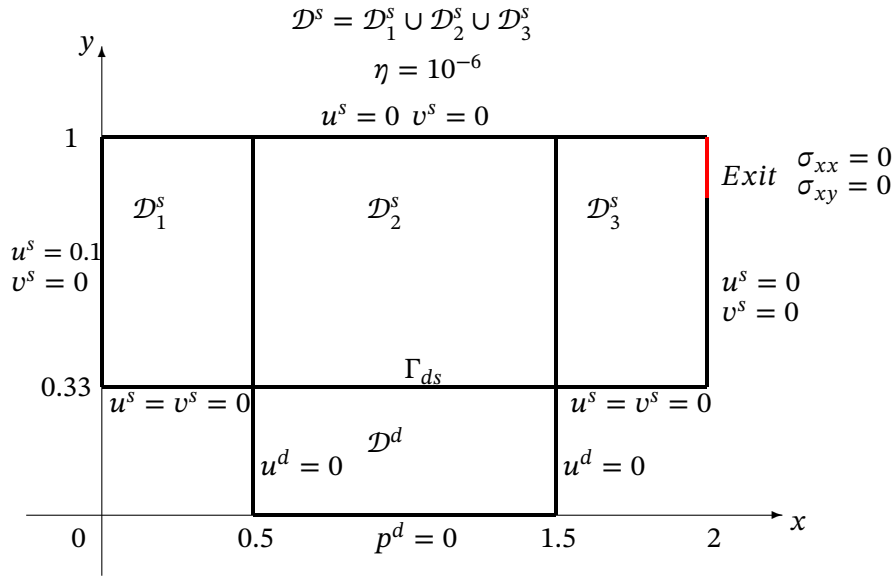
Figure 5.10: Geometry of the 4-block problem with BJS interface condition. Subdivision of the domain $\mathcal{D}$ into a free-flow subregion $\mathcal{D}^s$ and a porous subdomain $\mathcal{D}^d$ by an internal interface Γ_{ds}.

(a) $\lambda_{cx} = 2, \lambda_{cy} = 0.02$
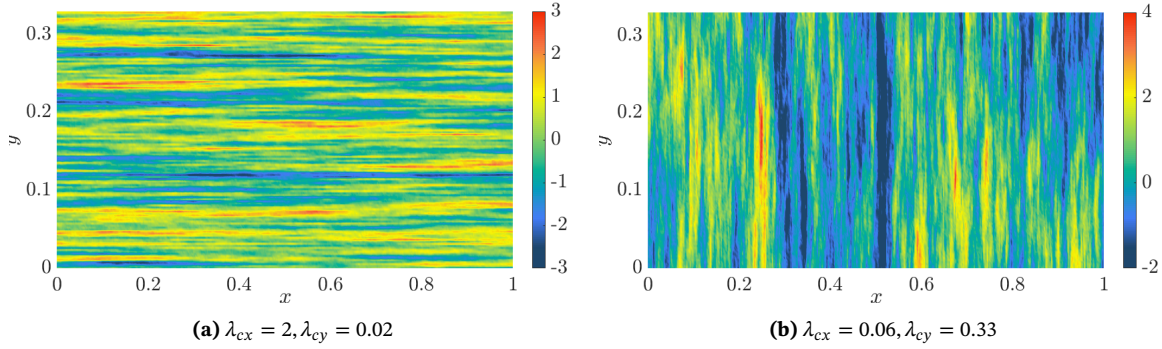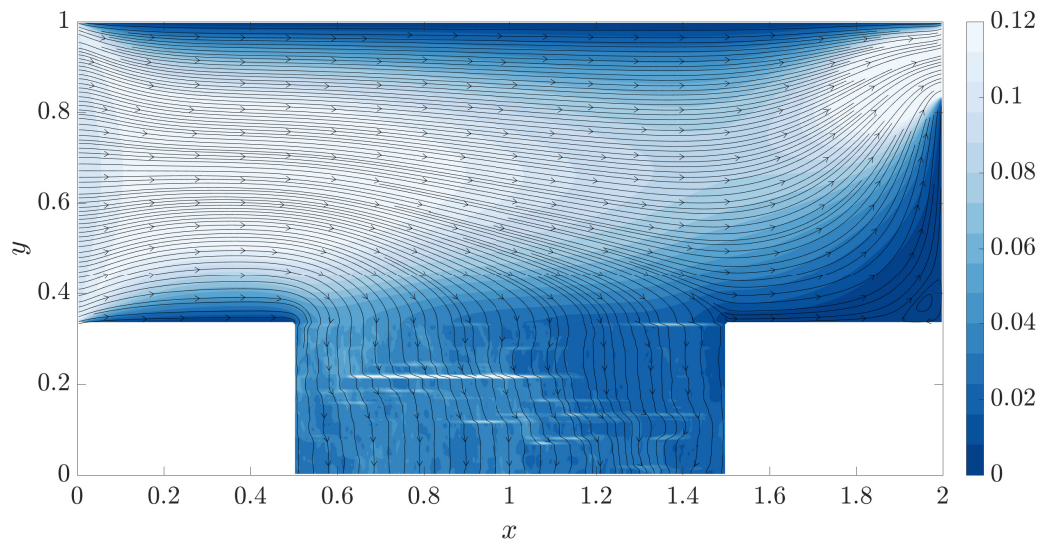
(b) $\lambda_{cx} = 0.06, \lambda_{cy} = 0.33$

Figure 5.11: Examples of Gaussian random fields ($\log k$) with (a) horizontal layers and (b) vertical layers. For both cases $\nu_c = 0.5$ and variance $\sigma_c^2 = 1$.

in (4.25):

$$\begin{cases} C_\Phi(\mathbf{x_1}, \mathbf{x_2}) = \sigma_c^2 \dfrac{2^{1-\nu_c}}{\Gamma(\nu_c)} \left(2\sqrt{\nu_c}\tilde{r}\right)^{\nu_c} k_{\nu_c}\left(2\sqrt{\nu_c}\tilde{r}\right), \\ \tilde{r} = \sqrt{\dfrac{(x_1 - x_2)^2}{\lambda_{cx}^2} + \dfrac{(y_1 - y_2)^2}{\lambda_{cy}^2}} \quad \text{with} \quad \mathbf{x_1} = (x_1, y_1), \mathbf{x_2} = (x_2, y_2). \end{cases} \tag{5.65}$$

Here the parameters have similar definitions as in (4.25). It is pointed out that the above Matérn model differs from the anisotropic model (4.25) only by the rotation parameter. In these experiments, we use horizontally and vertically layered Gaussian fields, see Figure 5.11. We also present the streamlines and magnitude ($\sqrt{u^2 + v^2}$) of the velocity fields generated from the horizontally (Figure 5.12a) and vertically (Figure 5.12b) layered permeabilities. These solutions are based on a staggered mesh with $h = 1/384$ such that the number of cells on different blocks are $\mathcal{D}_1^s = 192 \times 256$, $\mathcal{D}_2^s = 384 \times 256$, $\mathcal{D}_3^s = 192 \times 256$ and $\mathcal{D}^d = 384 \times 128$. The effect of layered permeability is clearly visible in both figures resulting in randomly layered high and low mobility zones. The flow intensity through the right exit depends on the particular realization of the random fields. In Figure 5.13, we show the mean $\pm$ standard deviation multigrid convergence for three parameter sets: isotropic, small vertical

(a) $\lambda_{cx} = 2, \lambda_{cy} = 0.02$
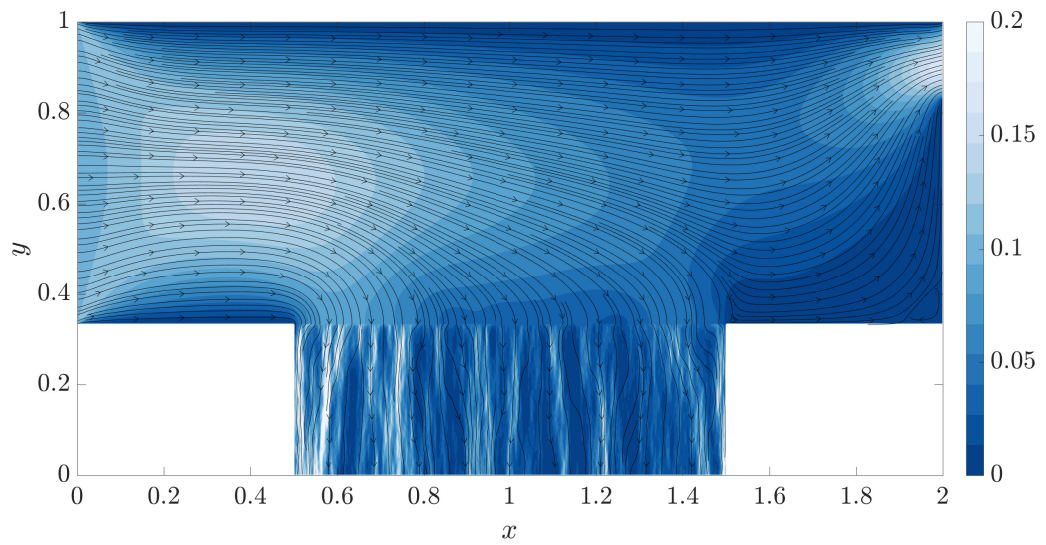
(b) $\lambda_{cx} = 0.06, \lambda_{cy} = 0.33$

Figure 5.12: Flow field generated for (a) horizontally and (b) vertically layered porous media. These flows are computed on a staggered mesh with $h = 1/384$.
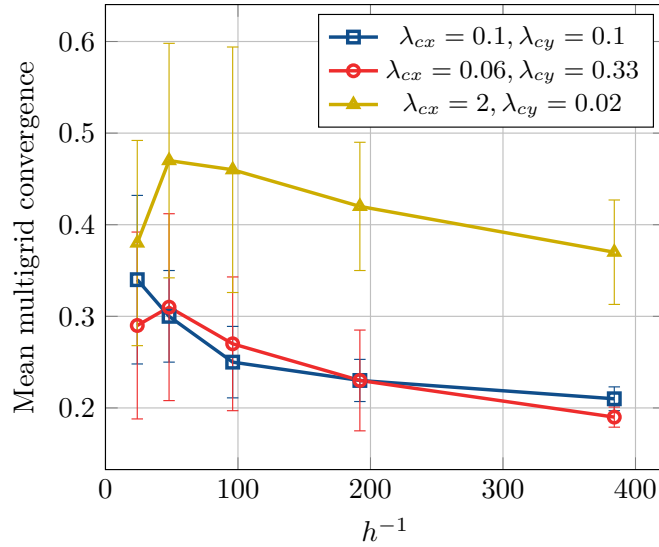
Figure 5.13: Mean multigrid convergence $\langle q \rangle$ along with the standard deviation for $W(2,2)$-cycle.

and small horizontal correlations. For all these cases, the smoothness is taken as $\nu_c = 0.5$ and variance $\sigma_c^2 = 1$. For the isotropic case with $\lambda_{cx} = 0.1, \lambda_{cy} = 0.1$, a similar convergence behaviour as observed for the no-slip interface condition test case is seen. For the vertically layered case with $\lambda_{cx} = 0.06, \lambda_{cy} = 0.33$, fast convergence rates are achieved. A deterioration in the multigrid convergence for horizontally layered permeability for $\lambda_{cx} = 2, \lambda_{cy} = 0.02$ is also observed. This is due to very high contrast between x-y correlation lengths ($\lambda_{cx}/\lambda_{cy} = 100$). The mean convergence rates improves with grid refinement for all these hyper-parameters. Multigrid with semi-coarsening or line-wise smoothers may further improve the convergence [14].

5.5.2 MLMC convergence

So far, we have benchmarked the performance of the multigrid solver and accuracy of the FV scheme for the random Darcy-Stokes flow. Now, we use the MLMC estimator to compute moments of the stochastic transport in the Darcy-Stokes system. Note that one can also apply the MLMC estimator together with the multigrid solver to efficiently compute the statistical moments of the random flow field. The implementation details of the plain and multilevel Monte Carlo methods can be found in Sections 2.2 and 2.3, respectively.

We consider the transport in the same 2-block problem with the no-slip condition (Section 5.5.1). The source term is taken to be zero, i.e., $f^t \equiv 0$. To investigate the convergence of the proposed spatio-temporal discretization for the fully coupled system, we use two deterministic inflow boundary conditions, based on a Gaussian plume (GP) and a discontinuous square-wave (SW), defined as

$$c_{in}(0, y, t) = \exp\left(-\frac{|y - 1.5|}{\lambda^2}\right), \quad \text{for} \quad \lambda = 0.1, y \in [1, 2], \quad t \geq 0, \tag{5.66}$$

$$c_{in}(0, y, t) = 1, \quad \text{for} \quad |y - 1.5| \leq 1/8, \quad t \geq 0. \tag{5.67}$$

respectively. The initial concentration c_0 is obtained from c_{in} at $t = 0$. Uncertainty in the inflow condition or initial concentration can be easily incorporated within this test case. The dispersion parameter in the Stokes region is taken as $D = 10^{-6}$ and for the Darcy region parameters in (5.21)
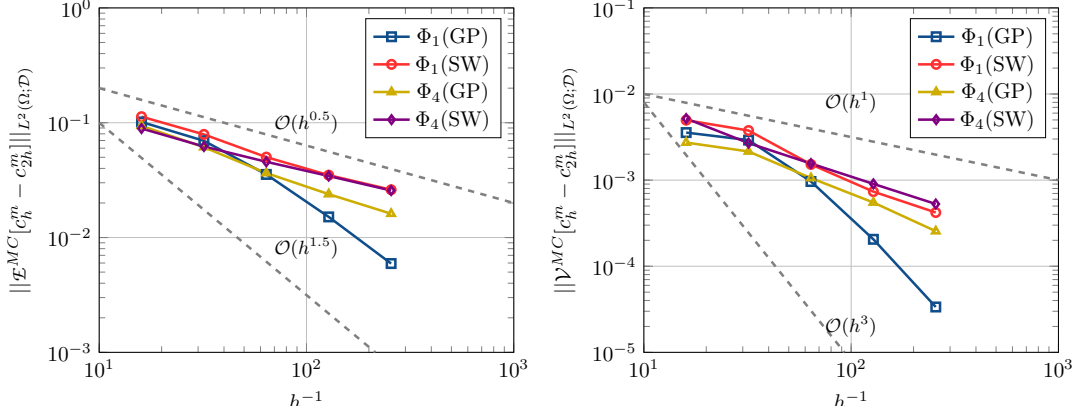
Figure 5.14: The convergence of the FV bias and the level-dependent variance for different combinations of Matérn parameter and inflow boundary conditions.

are $D_L = D_T = 10^{-5}$ with $D^* = 0$ resulting in a diagonal dispersion tensor. We consider the porosity $\phi = 1$ for the Stokes region and $\phi = 0.4$ for the porous medium.

In Figure 5.14, we show the convergence of the FV bias (left) along with the decay of the level dependent variance (right) for four combinations of Matérn parameters and inflow conditions. For the Gaussian plume model (5.66) the bias decays with the rate of $\mathcal{O}(h^{\nu_c})$ depending on the smoothness parameter of the permeability field as the dominant error comes from the discretization of the Darcy-Stokes velocity. As the proposed scheme is second-order accurate, we get $\mathcal{O}(h^2)$ for $\nu_c \geq 2$. For the discontinuous square-wave function, we observe a rate of $\mathcal{O}(h^{0.5})$ regardless of the smoothness parameter. Further, we observe that the variance decays at twice the rate of the FV bias. The empirical values of the MLMC parameters α and β can be derived from these plots. It is pointed out that for the easiest parameter set Φ_1(GP), we can get a cubic decay of samples with increasing level and for the hardest case Φ_4 (SW), we can achieve a linear decay only.

To illustrate the behaviour of the contaminant transport problem in the random flow field and compare the first-order UDS and the flux-limited QUICK scheme, we present two examples with the concentration distribution after $t = 6.00$ in Figure 5.15 where the velocity fields are generated using the Matérn parameter set Φ_3 (top) and Φ_4 (bottom), respectively with the discontinuous inflow (5.67). The contaminant plume remains compact as long as it remains in the free-flow region but as soon as it reaches the interface, it starts to spread out due to the heterogeneity of the porous medium. The QUICK method is clearly less diffusive than the UDS and the flux limiter is able to remove oscillations near sharp discontinuities. Also, the velocity profile and concentration distribution from Φ_3 are visibly more smooth compared to Φ_4.

To demonstrate the convergence of the MLMC method we will consider the most difficult case Φ_4 (SW). To measure the accuracy of the MLMC and MC estimators, we use the following L^2-based relative error measure

$$\varepsilon_{rel} := \frac{||\mathcal{E}_{ref}[c(\cdot, t^m)] - \mathcal{E}[c_h^m]||_{L^2(\mathcal{D}_h)}}{||\mathcal{E}_{ref}[c(\cdot, t^m)]||_{L^2(\mathcal{D}_h)}}, \tag{5.68}$$

where $\mathcal{E}$ can be replaced with either $\mathcal{E}_{h_L}^{ML}$ (MLMC) or $\mathcal{E}_N^{MC}$ (MC). The reference solution is computed on a relatively fine grid. For the current test problem, the reference solution is presented in Figure 5.16, showing the mean and variance of the spatial concentration distribution at $t = 6.00$. The samples required for the MLMC estimator were computed using the multigrid solver with $W(2, 2)$-cycle to obtain the velocity field and the ADI time-stepping was applied to solve the transport equation. We observe that the plume remains compact in the free-flow region with very insignificant variance. In the Darcy region, the spatial mean and variance are symmetrically distributed. The
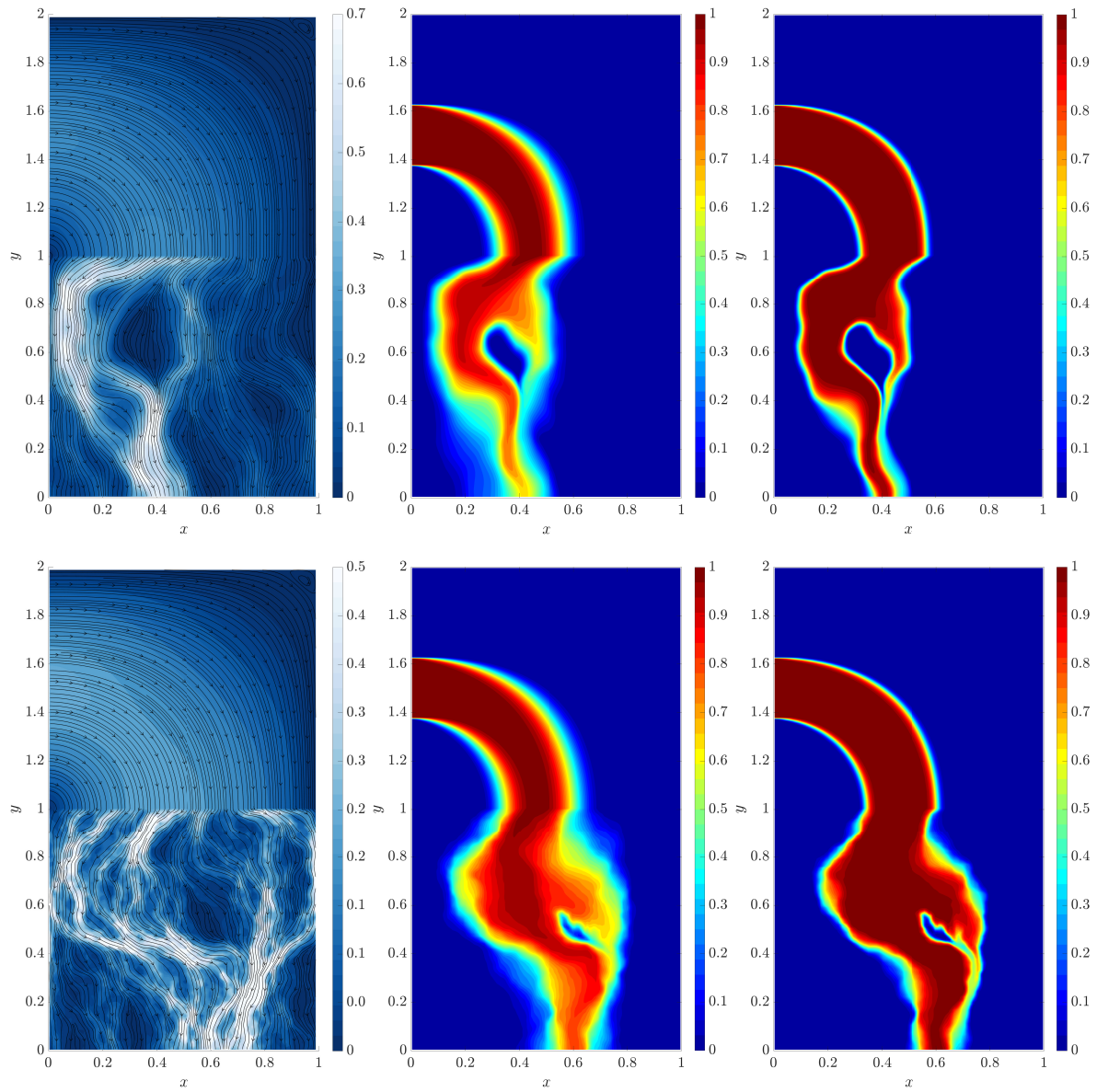
Figure 5.15: Transport solution at $t = 6.00$ from UDS (middle) and flux-limited QUICK (right) scheme for same velocity field (left) with $h = \Delta t = 1/128$. The permeability field is generated using Φ_3 (top) and Φ_4 (bottom).

(a) Mean of $c(x, y, t = 6.00)$ **(b)** Variance of $c(x, y, t = 6.00)$
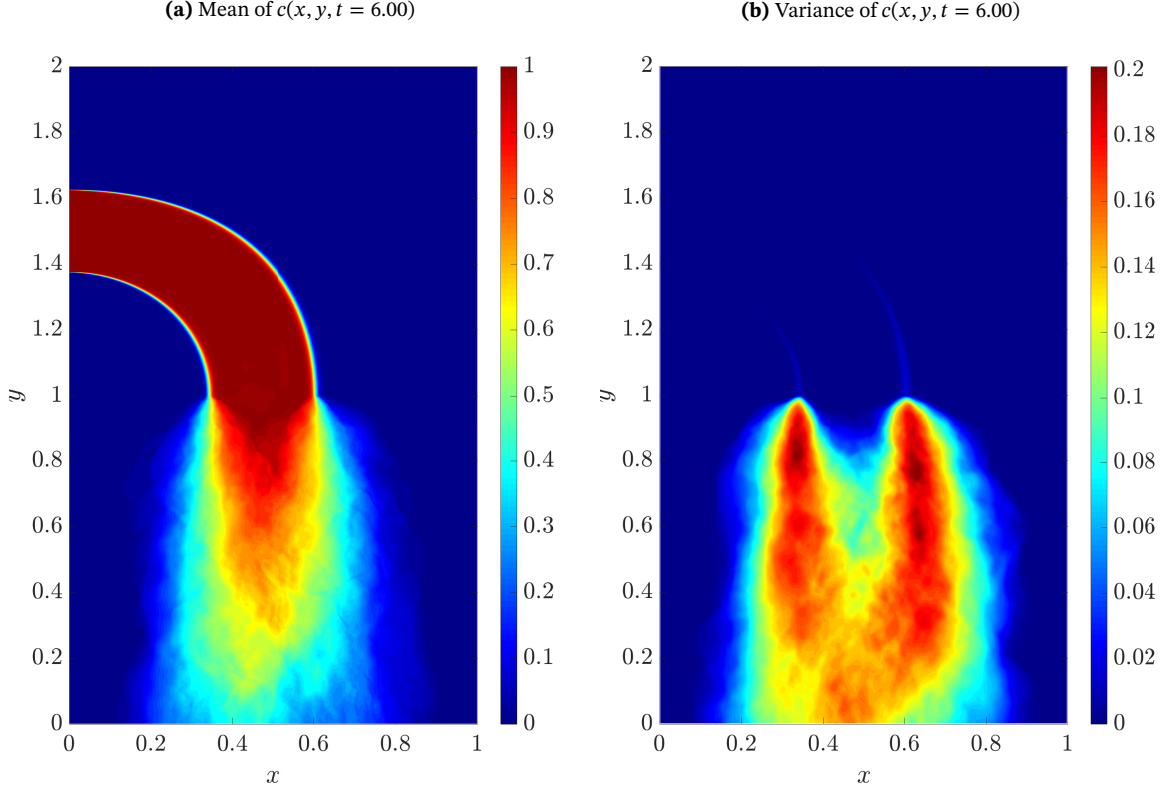

Figure 5.16: Statistical solution of uncertain transport problem with parameter Φ_4 at $t = 6.00$ computed using the MLMC estimator. High variance is observed in the Darcy region. Sharp gradients are preserved in the Stokes domain with very insignificant variance.

reference solution is computed using a six level MLMC estimator ($L = 6$) with the coarsest level $h_0 = \Delta t_0 = 1/16$ and the finest level $h_L = \Delta t_L = 1/512$. We use the fixed sampling strategy discussed in Section 2.3.3 with $N_L = 8$ and the number of samples on coarser levels is derived using relation $N_\ell = \lceil N_L 2^{\beta(L-\ell)} \rceil$. For case Φ_4 (SW), we get $\alpha \approx 0.5$ and $\beta \approx 2\alpha \approx 1$ (see Figure 5.14) resulting in a sample sequence $N_\ell = N_L 2^{(L-\ell)}$. Recall that with these rates, we end up with $\gamma > \beta$ resulting in an optimal MLMC estimator, see Eq. (2.27) from Chapter 2. For comparison, we also implement the standard MC estimator where the number of samples is derived by balancing the sampling error with FV error, i.e. $N = \mathcal{O}(h^{-2\alpha}) = \mathcal{O}(h^{-1})$. We chose $N = 8$ for grid $h = 1/16$ and double the number of samples with subsequent refinements, when using the plain MC method.

In Figure 5.17 (left), we compare the accuracy of the MLMC and the standard MC estimators. For each h, we repeated the experiment eight times and the average relative error denoted by $\bar{\varepsilon}_{rel}$ is plotted. As expected, both the MLMC and MC estimators were able to achieve a similar accuracy converging as $\mathcal{O}(h^{0.5})$. Also, in Figure 5.17 (right), we show the cost scaling for both estimators. The CPU time for the MLMC estimator grows with a slower rate than the standard MC estimator and we observe a speed-up of more than an order of magnitude for $h = 1/256$. For reference, we also indicate the theoretical cost (dotted grey lines) for the MC estimator ($\mathcal{W}^{-1/8}$) and the MLMC estimator ($\mathcal{W}^{-1/6}$), when using an optimal solver (or $\gamma = 3$). Here, we have only presented the results corresponding to the most difficult test case Φ_4 (SW). In general, the asymptotic cost of the MLMC method further improves with the smoothness of the random field as well as with smooth inflow conditions. For example, it is possible to obtain a complexity of about $\mathcal{W}^{-1/2}$ (ignoring the logarithmic term) for case Φ_1 (GP) as $\beta \approx \gamma \approx 3$.

We also remark that the proposed multigrid MLMC method can be extended to parallel architectures. Specifically, one can achieve three degrees of parallelization over samples, levels and
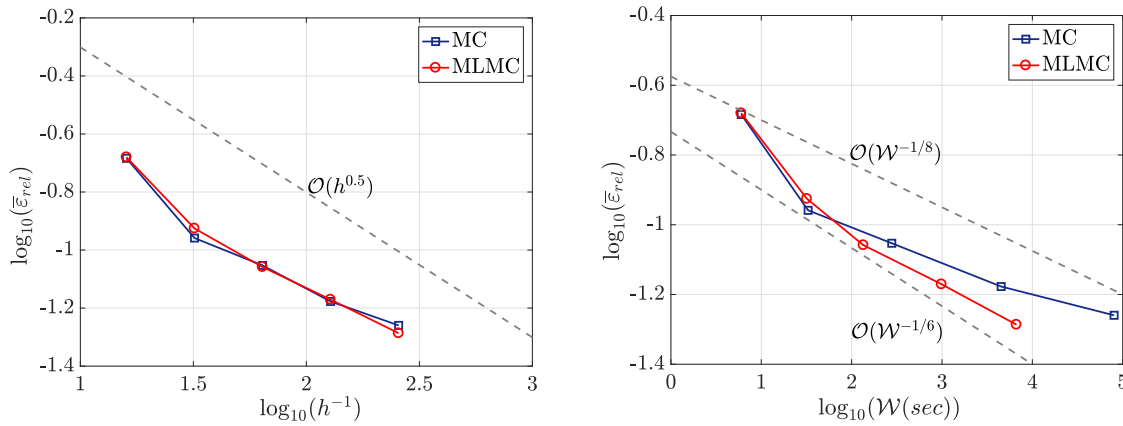
Figure 5.17: Comparison of relative error convergence $\bar{\varepsilon}_{rel}$ with grid size (left) and computational work $\mathcal{W}$ versus accuracy (right) for MLMC and MC estimators.

the multigrid solver, see [31, 113, 114] for recent advances. Out of these, the most challenging part is the parallelization over the solver that should be considered only on finer levels when the communication overhead becomes negligible. Moreover, the grid partitioning method explored in this chapter can be very effective as each partitioned block can be relaxed on separate processors. We have already demonstrated that it is possible to obtain a satisfactory multigrid convergence using the 4-block example earlier.

5.6. Conclusions

In this chapter, we described an MLMC algorithm for uncertainty quantification of advection-dominated transport in a coupled Darcy-Stokes system with uncertain permeability. An important contribution of this work is the robust multigrid solver that can solve the coupled Darcy-Stokes flow with highly heterogeneous permeability field very efficiently. The new monolithic multigrid method achieves *textbook multigrid convergence* for a wide range of Matérn parameters, thus making it a highly suitable solver for MLMC applications. We combined a flux-limited QUICK scheme with an ADI time-stepping resulting in a second-order accurate spatio-temporal discretization of the stochastic transport equation that can be solved optimally. This implicit version also resulted in a scheme that is stable on very coarse grids and greatly helped in reducing the cost of the MLMC method. Lastly, we showed that for very rough problems, we can attain an asymptotically optimal MLMC estimator that has the same computational complexity as its deterministic counterpart.

<div style="text-align: right; font-size: 3em;">6</div>

Variably saturated flow in a porous media

This chapter[1] presents a multilevel Monte Carlo (MLMC) method for the uncertainty quantification of variably saturated porous media flow that are modeled using the Richards' equation. We propose a stochastic extension for the empirical models that are typically employed to close the Richards' equations. This is achieved by treating the soil parameters in these models as spatially correlated random fields with appropriately defined marginal distributions. As some of these parameters can only take values in a specific range, non-Gaussian models are utilized. The randomness in these parameters may result in path-wise highly nonlinear systems so that a robust solver with respect to the random input is required. For this purpose, a solution method based on a combination of the modified Picard iteration and a cell-centered multigrid method for heterogeneous diffusion coefficients is utilized. Moreover, we propose a non-standard MLMC estimator to solve the resulting high-dimensional stochastic Richards' equation. The improved efficiency of this multilevel estimator is achieved by parametric continuation that allows us to incorporate simpler nonlinear problems on coarser levels for variance reduction and the original strongly nonlinear problem is solved only on the finest level. Several numerical experiments are presented showing computational savings obtained by the new estimator compared to the standard MLMC estimator.

6.1. Introduction

Mass transport through a variably saturated porous medium can be accurately predicted using the Richards' equation [21]. This modeling approach is of critical importance for several physics and engineering problems, for instance, when studying aquifer recharge via rainfall infiltration, or for understanding the environmental impact of mining operations. When reliable measurements of the hydraulic properties are available, numerical solutions originating from the Richards' equation have been reasonably successful for transport prediction in a broad range of soil types.

Different formulations for the Richards' equation are available in the literature, along with well established mathematical theory, such as the pressure head, the water content or a mixed formulation, see e.g. [115–120]. The aforementioned formulations contain nonlinearities due to a parametric dependence of the pressure head on the saturation and the relative hydraulic conductivity. Depending on the soil parameters, these nonlinearities range from mild to strong. The extreme

---

[1]Note that the notations in this chapter are slightly different from the previous chapters, as we follow the notations commonly used in the literature related to variably saturated flow.

This chapter is based on the article "A parametric acceleration of multilevel Monte Carlo convergence for nonlinear variably saturated flow.", *Preprint arXiv:1903.08741*, (2019) [26]. (Submitted for publication)

sensitivity of the soil parameters on the output necessitates accurate measurements of the hydraulic properties. For many realistic problems, complete information of these quantities is however not available. In such scenarios, these parameters may be modeled in a probabilistic framework and the solution output may be expressed by means of a prediction interval (with mean and variance), rather than as a single value. Such approaches are nowadays common in the case of saturated groundwater flow, where uncertainties are included when modeling the hydraulic conductivity as a spatially correlated lognormal random field [18–20]. The purpose of the present work is to develop and analyze a stochastic extension of the Richards' equation, along with an efficient numerical method to solve the resulting nonlinear partial differential equation with random coefficients.

Previous work on the uncertainty quantification (UQ) of unsaturated flows was often based on an uncertain hydraulic conductivity [121–124]. In addition to that, in the present work, we introduce stochasticity in the so-called van Genuchten and Mualem model [125, 126], which is typically utilized to close the Richards' equation. This model provides a closed-form analytic expression for the unsaturated hydraulic conductivity based on a *sigmoid* type function for the soil-water retention curve. This curve is defined by four independent parameters that are estimated by curve-fitting, based on field measurements. Typically, these parameters are fixed throughout the domain during numerical simulations, assuming the soil to be homogeneous. Realistic models should however also incorporate the intrinsic heterogeneity in the soil. Therefore, we model these soil parameters as random variables with a certain, specified probability distribution and spatial correlations. To assure the well-posedness of the Richards' equation, these parameters should be within a certain range. Thus, the probability distributions for these parameters are chosen such that the random samples will be in the domain of validity for these parameters. A practical choice is to employ non-Gaussian random fields with marginal distributions from expert knowledge or from field measurements.

With the stochastic Richards' equation formulated, an appropriate UQ technique is required to compute the statistical moments of the desired quantities of interest (QoIs). This choice primarily depends on the number of uncertainty dimensions. Other practical factors, such as ease of implementation and availability of an iterative solver which is robust with respect to the random input, also play a role in the selection of a suitable UQ technique. The proposed stochastic extension of the Richards' equation results in a very high-dimensional problem, and the use of deterministic sampling approaches such as polynomial chaos expansion, stochastic collocation or stochastic Galerkin is therefore limited. For these UQ methods, the cost grows exponentially with the number of random inputs. Furthermore, a deterministic sampling approach may not adequately represent those regions in the stochastic space where strong nonlinearity may be encountered. In previous works of Zhang [123, 127], the moments method was applied for the uncertainty quantification of solutions of the Richards' equation. The main disadvantage of a moment-based method is that it can only be reliably employed when the effect of uncertain inputs is mild and largely linear. For the proposed stochastic formulation of the Richards' equation, MC based sampling approaches are the *methods of choice*, due to their dimension independent convergence. Moreover, MC-type methods due to the randomness enables better exploration of the stochastic space − a quality which is not usually shared by the deterministic approaches.

The plain MC method can be highly expensive for the stochastic Richards' equation as cost per sample is very high. The original, grid-based MLMC estimator may be utilized to solve the stochastic Richards' equation, however, this approach may not be the most efficient, especially not when strongly nonlinear problems need to be solved. Such problems require a very fine spatio-temporal mesh thereby restricting the use of coarse grids to improve the efficiency of the MLMC estimator. In this work, we utilize a non-standard MLMC estimator based on the *parametric continuation* technique. Continuation methods for solving nonlinear PDEs are very popular in engineering applications [128–133]. Within continuation methods, a nonlinearity dictating parameter Θ is

introduced in the interval $\Theta_0 \leq \Theta \leq \Theta_*$ where the solution $p(\Theta_0)$ corresponds to a linear (or mildly nonlinear) problem and $p(\Theta_*)$ to the target strongly nonlinear. The key idea is to march from $p(\Theta_0)$ to $p(\Theta_*)$ in small steps of size $\delta\Theta$, where at each step we use the solution from the previous step as an initial guess. Usually, Θ is some physical parameter, for e.g. the Reynolds number, the Mach number, etc. In the current work, we use parametric continuation to obtain variance reduction within the multilevel Monte Carlo framework. This is achieved by solving simpler nonlinear problems on coarser levels and the target strongly nonlinear problem is only solved on the finest level. This new estimator allows us to incorporate comparatively coarser spatio-temporal grids in the MLMC hierarchy and, as such, the computational cost of each estimator in the telescopic sum is greatly reduced.

We furthermore propose a solution method for Richards' equation based on a combination of the modified Picard method [115] and a cell-centered multigrid, as proposed in [24] (also see Chapter 4). We benchmark the performance of this combined solver in a probabilistic framework. A number of tests for a wide range of soil parameters and for hydraulic conductivities with different heterogeneity levels are performed.

The rest of the chapter is organized as follows. In Section 6.2, we briefly discuss the deterministic Richards' equation along with the van Genuchten-Mualem parameterization. Section 6.3 describes the stochastic Richards' equation as well as the modeling of various uncertain soil parameters. The description of the modified Picard method in combination with the cell-centered multigrid method is provided in Section 6.4. Also, in this section, we present some numerical experiments to assess the performance of the combined solver for an infiltration problem. The non-standard MLMC estimator is explained in Section 6.5 and its performance is analysed in Section 6.6. Finally, some conclusions are drawn in Section 6.7.

6.2. Deterministic Richards' equation

We begin by describing the deterministic version of the problem. The governing equations are defined in a bounded domain $\mathcal{D} \subset \mathbb{R}^n$, with the boundary $\partial\mathcal{D}$ and a finite time interval $\mathcal{T} = (0, T_{final}]$, with $T_{final} < \infty$. The classical Richards' equation is a result of coupling the mass conservation equation of the water-phase and the Darcian flow, i.e.,

$$\phi\frac{\partial S_w}{\partial t} + \nabla \cdot \mathbf{q} = f \quad \text{in} \quad \mathcal{D} \times \mathcal{T}, \tag{6.1}$$

$$\mathbf{q} = -K_s K_{rw}(\nabla p + z) \quad \text{in} \quad \mathcal{D} \times \mathcal{T}, \tag{6.2}$$

respectively, subject to boundary and initial conditions:

$$p = p_0 \quad \text{in} \quad \mathcal{D}, \quad t = 0, \tag{6.3}$$

$$p = g_D \quad \text{in} \quad \Gamma_D \times \mathcal{T}, \tag{6.4}$$

$$\mathbf{q} \cdot \mathbf{n} = g_N \quad \text{in} \quad \Gamma_N \times \mathcal{T}, \tag{6.5}$$

where $\phi[L^3/L^3]$ is the porosity; $S_w[L^3/L^3]$ is the water-phase saturation; $\mathbf{q}$ is the Darcy flux, which depends on the pressure head, $p[L]$, and the depth $z[L]$ in the vertical direction; $K_s[L/T]$ represents the saturated hydraulic conductivity field at saturation; K_{rw} is the relative conductivity of the water phase with respect to air and f is the source/sink term. The initial pressure head value is given by p_0. The quantities g_D and g_N denote, respectively, the Dirichlet and Neumann boundary conditions that are imposed at the boundaries Γ_D and Γ_N, respectively, with $\mathbf{n}$ the unit normal vector to Γ_N.

The coupling of (6.1) and (6.2) may result in different variants of the Richards' equation, such as the pressure head, the moisture content and the mixed formulation. The mixed formulation of the Richards' equation is given by:

$$\frac{\partial\theta(p)}{\partial t} - \nabla \cdot (K_s K_{rw}(\nabla p + z)) = f \quad \text{in} \quad \mathcal{D} \times \mathcal{T}. \tag{6.6}$$

It is obtained by substituting the moisture content, i.e., $\theta = \phi S_w(p)$. By using

$$\frac{\partial \theta(p)}{\partial t} = C(p)\frac{\partial p}{\partial t},$$

the above PDE can be reformulated into the pressure head formulation:

$$C(p)\frac{\partial p}{\partial t} - \nabla \cdot (K_s K_{rw}(\nabla p + z)) = f \quad \text{in} \quad \mathcal{D} \times \mathcal{T}, \tag{6.7}$$

where $C(p) = \frac{\partial \theta}{\partial p}$ is the specific moisture capacity. It is well-known that numerical solutions originating from the pressure head formulation may give rise to a significant mass balance error, resulting in an inaccurate prediction of the infiltration depth.

Numerical methods based on the mixed form (using finite differences or mass-lumped finite elements) are popular as they result in mass conservation schemes [115]. Therefore, we will work with the mixed form (6.6) of the Richards' equation.

6.2.1   Van Genuchten-Mualem model

To complete the PDE formulations, (6.6) or (6.7), closure models for approximating K_{rw} and θ are required. A number of models have been presented in the literature and the most popular ones are by Brooks-Corey [134] and van Genuchten-Mualem [125, 126]. These two models employ nonlinear constitutive relations for K_{rw} and p, and for θ and p, respectively. We consider the parameterization introduced by van Genuchten and Mualem here. For the saturation, van Genuchten [125] proposed the following analytic formula:

$$S_w(p) = \frac{\theta(p) - \theta_r}{\theta_s - \theta_r} = \begin{cases} (1 + (|\alpha p|)^n)^{-m}, & p < 0, \\ 1, & p \geq 0, \end{cases} \tag{6.8}$$

where θ_s and θ_r are the saturated and residual water contents, respectively, and $\alpha[L^{-1}]$, n and $m = 1 - n^{-1}$ are obtained by fitting data characterizing the statistics of the soil. Specifically, the parameter α provides a measure of the average pore-size in the soil matrix and n is related to the pore-size distribution of the soil [135].

We may derive the specific moisture content, $C(p)$, analytically from (6.8), as

$$C(p) = \begin{cases} (\theta_s - \theta_r)\alpha mn(1 + |\alpha p|^n)^{-(m+1)}|\alpha p|^{n-1}, & p < 0, \\ 0, & p \geq 0. \end{cases} \tag{6.9}$$

In previous work, Mualem [126] derived a closed-form expression for K_{rw}, which is given by:

$$K_{rw} = S_w^{1/2}\left[\int_0^{S_w} dS_w/p \bigg/ \int_0^1 dS_w/p\right]^2. \tag{6.10}$$

Using (6.8), the above integral equation reduces to the following analytic expression:

$$K_{rw}(p) = \begin{cases} S_w(p)^{1/2}\left(1 - (1 - S_w(p)^{1/m})^m\right)^2, & p < 0, \\ 1, & p \geq 0. \end{cases} \tag{6.11}$$

The complexity of the numerical solution of the Richards' equation depends on the values of the parameters n and α. For $n \in (1, 2)$ and $p \to 0$, the relative hydraulic conductivity $K_{rw}(p)$ is not

Lipschitz continuous and the derivative $K'_{rw}(p)$ becomes infinite as p approaches zero [135, 136]. Moreover, for small values of n, a sharp K_{rw} vs. p profile is encountered. Similarly, for large values of the parameter α, the pressure head exhibits a transition behaviour with a steep gradient from the saturated to the unsaturated region. In general, for a small n or for large α strong nonlinearities are encountered, thus implying convergence issues for nonlinear iterative techniques such as the Newton or Picard methods.

6.3. Stochastic Richards' model

Here, we describe a stochastic extension of the van Genuchten model. We assume that the unknown soil parameters belong to the probability space $(\Omega, \mathbb{F}, \mathbb{P})$, where Ω is the sample space with σ-field $\mathbb{F} \subset 2^{\Omega}$ as a set of events and the probability measure $\mathbb{P} : \mathbb{F} \to [0, 1]$.

The stochastic extension is based on modeling the soil parameters as spatially correlated random fields in order to incorporate spatial heterogeneity. For the saturated hydraulic conductivity, K_s, it is standard practice to model it as a lognormal random field, as follows,

$$K_s(\mathbf{x}, \omega) = K_s^{(bl)}(\mathbf{x}) \exp(Z(\mathbf{x}, \omega)), \qquad \mathbf{x} \in \mathcal{D}, \omega \in \Omega, \tag{6.12}$$

where $K_s^{(bl)}(\mathbf{x})$ is the baseline hydraulic conductivity and $Z(\mathbf{x}, \omega)$ is a zero mean Gaussian random field with a specified covariance kernel. So,

$$\mathbb{E}[Z(\mathbf{x}, \cdot)] = 0, \tag{6.13}$$
$$\text{Cov}(Z(\mathbf{x_1}, \cdot), Z(\mathbf{x_2}, \cdot)) = \mathbb{E}[Z(\mathbf{x_1}, \cdot)Z(\mathbf{x_2}, \cdot)], \quad \mathbf{x_1}, \mathbf{x_2} \in \mathcal{D}. \tag{6.14}$$

We consider an anisotropic Matérn covariance function (see also Chapter 5) , C_Φ, defined as

$$\begin{cases} C_\Phi(\mathbf{x_1}, \mathbf{x_2}) = \sigma_c^2 \dfrac{2^{1-\nu_c}}{\Gamma(\nu_c)} \left(2\sqrt{\nu_c}\tilde{r}\right)^{\nu_c} K_{\nu_c}\left(2\sqrt{\nu_c}\tilde{r}\right), \\[2ex] \tilde{r} = \sqrt{\dfrac{(x_1 - x_2)^2}{\lambda_{cx}^2} + \dfrac{(z_1 - z_2)^2}{\lambda_{cz}^2}}, \quad \text{with} \quad \mathbf{x_1} = (x_1, z_1), \mathbf{x_2} = (x_2, z_2). \end{cases} \tag{6.15}$$

Here, we denote the gamma function by Γ and by K_{ν_c} the modified Bessel function of the second kind. The Matérn function is characterized by the parameter set $\Phi = \{\nu_c, \lambda_{cx}, \lambda_{cz}, \sigma_c^2\}$ where ν_c is the smoothness, λ_{cx} and λ_{cz} are correlation lengths along x- and z-directions, respectively, and σ_c^2 is the marginal variance. The sampling procedure for Gaussian random fields is provided in Section 3.6.

6.3.1 Sampling of non-Gaussian random fields

For sampling the van Genuchten parameters, $\alpha(\mathbf{x}, \omega), n(\mathbf{x}, \omega), \theta_s(\mathbf{x}, \omega), \theta_r(\mathbf{x}, \omega)$ in Section 6.2.1, we employ random fields with *non-Gaussian marginal* distributions. This choice of distributions is practical as these parameters can only take values in a certain range, see e.g. [136]. We introduce stochasticity in the parameters via an additive noise,

$$\alpha(\mathbf{x}, \omega) = \alpha^{(bl)}(\mathbf{x}) + \varepsilon_\alpha(\mathbf{x}, \omega), \tag{6.16}$$

where $\alpha^{(bl)}(\mathbf{x})$ is the deterministic baseline value and $\varepsilon_\alpha(\mathbf{x}, \omega)$ is a random field with a non-Gaussian marginal distribution and covariance C_Φ. Notations are analogously for the other three van Genuchten parameters. Next, we describe a technique proposed in [137] for the point-wise transformation of a standard Gaussian random field to a non-Gaussian random field.

Non-Gaussian random fields are difficult to simulate as they are not uniquely determined by their mean and variance. There are however different techniques available for simulating non-Gaussian fields, see e.g. [137, 138]. In this work, we will follow a basic approach based on a

generalized Polynomial Chaos expansion [137], which approximates the non-Gaussian field in terms of a weighted combination of Hermite orthogonal polynomials of the standard Gaussian field,

$$Y(\mathbf{x}, \omega) \approx \sum_{j=0}^{N_{PC}} w_j \mathcal{H}_j(Z(\mathbf{x}, \omega)), \tag{6.17}$$

where $Y(\mathbf{x}, \omega)$ is the non-Gaussian random field (with a marginal distribution, e.g. the uniform distribution, gamma distribution, truncated normal, etc). $\mathcal{H}_j(Z)$ is the Hermite polynomial in Z of order j with weight w_j and N_{PC} is the order of the expansion. Hermite polynomials can be expressed as:

$$\mathcal{H}_0(Z) = 1, \quad \mathcal{H}_j(Z) = (-1)^j \exp(Z^2/2) \frac{d^j}{dx^j} \exp(Z^2/2), \quad j \in \mathbb{N}. \tag{6.18}$$

As Hermite polynomials are orthogonal with respect to the Gaussian measure, the weights can be evaluated using

$$w_j = \frac{\mathbb{E}[Y \mathcal{H}_j(Z)]}{\mathbb{E}[\mathcal{H}_j(Z)^2]}. \tag{6.19}$$

Here, the denominator is basically an expectation of a polynomial of the Gaussian random variable, which has an analytic expression. As the dependence between Y and Z is unknown, the expectation in the numerator is not explicit. Since the cumulative distribution for Y, defined as $F_Y(y) = \mathbb{P}\text{rob}(Y \leq y)$, is however known, one can utilize the relation $Y = F_Y^{-1}(F_Z(Z))$ to reformulate (6.19) as

$$w_j = \frac{1}{\mathbb{E}[\mathcal{H}_j(Z)^2]} \int_{I_Z} F_Y^{-1}[F_Z(z)] \mathcal{H}_j(z) dF_Z(z), \tag{6.20}$$

where I_Z is the support of Z in the range $(-\infty, \infty)$ and F_Y^{-1} representing the inverse of the distribution F_Y. Similarly, $F_Z(z) = \mathbb{P}\text{rob}(Z \leq z)$ is the cumulative distribution for standard Gaussian random variable Z. The one-dimensional integral (6.20) can be numerically computed using any conventional integration technique. For simplicity, we will use the Monte Carlo quadrature to approximate the above integral in our numerical experiments. These weights only need to be computed once, so that the cost of sampling a non-Gaussian random field with a stationary covariance function is of the same order as that of a Gaussian random field.

We will experiment here with both isotropic and anisotropic Matérn covariance models. In Table 6.1, the two Matérn parameters sets are listed, Φ_1 corresponding to an isotropic model and Φ_2 to an anisotropic model. In Figure 6.1, we present some samples of the random fields with a Gaussian and a uniform marginal distribution, for the two Matérn parameters. We use $N_{PC} = 6$ in (6.17) for generating the random fields with uniform marginal distribution. Due to a small correlation length and low spatial regularity, the numerical solutions of the PDE with random coefficients based on Φ_2 are comparatively more expensive to compute than those obtained with Φ_1. A comprehensive study on the computational cost of solving elliptic PDEs with different Matérn parameters can be found in [24] and Chapter 4. We will study the effect of covariance functions on the performance of the solver for the Richards' equation.

Φ_1	Φ_2
$(1.0, 0.2, 0.2, 1)$	$(0.5, 0.1, 0.01, 1)$

Table 6.1: Two combinations of Matérn parameters $\Phi = (\nu_c, \lambda_{cx}, \lambda_{cz}, \sigma_c^2)$ corresponding to isotropic (Φ_1) and anisotropic (Φ_2) random fields.
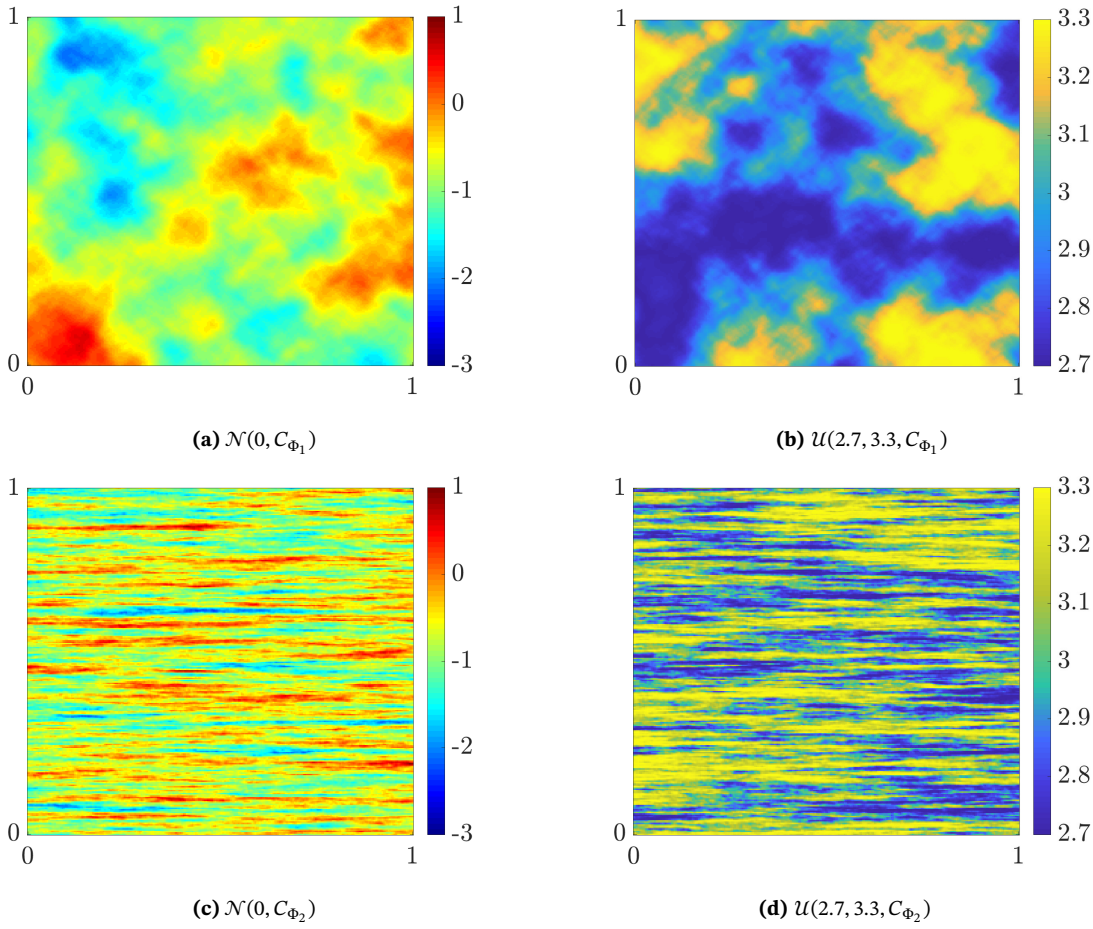
Figure 6.1: Samples of random fields generated using the isotropic Matérn parameter Φ_1 (top row) and the anisotropic parameter Φ_2 (bottom row); and standard normal marginal distribution (left column) and uniform marginal distribution (right column). The notation $\mathcal{U}(2.7, 3.3, C_\Phi)$ represents a random field with uniform marginal distribution, $\mathcal{U}[2.7, 3.3]$, with spatial correlation defined by C_Φ.

6.4. Modified Picard iteration combined with the cell-centered multigrid method

Algorithms based on the modified Picard iteration from Celia et al. [115] are often employed as efficient iterative solution methods for the Richards' equation. These methods are relatively easy to implement, as they do not require the computation of Jacobians and they also have low storage requirements. Within each modified Picard iteration, a diffusion equation with variable coefficients needs to be solved. For this, we propose to utilize the *cell-centered multigrid* (CCMG) for heterogeneous diffusion coefficients, as proposed in [24, 68, 81] (also see Section 4.2.2). The CCMG algorithm is efficient as it is constructed with a simple set of transfer operators and it has been demonstrated to perform well for a large class of highly heterogeneous and also jumping diffusion coefficients [24].

6.4.1 Modified Picard iteration

We briefly recall the fully-implicit Picard iteration for the mixed formulation of the Richards' equation from [115]. With Δt the time-step and for any integer $J > 1$, we define a uniform temporal grid by $\{t^j = j\Delta t, j = 0, \dots, J\}$. The iteration number within a time-step is denoted by an integer $k > 0$. For simplicity, we use a simplified notation for $\theta^{j,k} = \theta(p^{j,k})$ and $K^{j,k} = K_s K_{rw}(p^{j,k})$. The

backward Euler approximation of (6.6) is then written as

$$\frac{\theta^{j+1,k+1} - \theta^j}{\Delta t} - \nabla \cdot K^{j+1,k} \nabla p^{j+1,k+1} - \frac{\partial K}{\partial z}^{j+1,k} = f^{j+1}. \tag{6.21}$$

The key idea of the modified Picard iteration is the use of a Taylor expansion for $\theta^{j+1,k+1}$ with respect to p, i.e.

$$\theta^{j+1,k+1} = \theta^{j+1,k} + \frac{\partial \theta}{\partial p}^{j+1,k} (p^{j+1,k+1} - p^{j+1,k}) + \mathcal{O}\left(\delta p^2\right), \tag{6.22}$$

where the derivative $\frac{\partial \theta(p)}{\partial p} = C(p)$ is analytically computed by using (6.9). By neglecting the higher-order terms in (6.22) and substitution in (6.21), we get

$$C(p^{j+1,k})\frac{\delta p^{j+1,k}}{\Delta t} + \frac{\theta^{j+1,k} - \theta^j}{\Delta t} - \nabla \cdot K^{j+1,k} \nabla p^{j+1,k+1} - \frac{\partial K}{\partial z}^{j+1,k} = f^{j+1}, \tag{6.23}$$

with $\delta p^{j+1,k} = p^{j+1,k+1} - p^{j+1,k}$. The above equation can be expressed in the form:

$$C(p^{j+1,k})\frac{\delta p^{j+1,k}}{\Delta t} - \nabla \cdot K^{j+1,k} \nabla \delta p^{j+1,k} = \nabla \cdot K^{j+1,k} \nabla p^{j+1,k} + \frac{\partial K}{\partial z}^{j+1,k} + f^{j+1} - \frac{\theta^{j+1,k} - \theta^j}{\Delta t}. \tag{6.24}$$

The next pressure head iterate is obtained by the update $p^{j+1,k+1} = p^{j+1,k} + \delta p^{j+1,k}$. Notice that the left-hand side of the above equation is the residual associated with the Picard iteration, which should be equal to zero for a converged solution. Therefore, one may use $||\delta p^{j+1,k}||_\infty < \varepsilon_{PI}$ as a stopping criterion with $\varepsilon_{PI} > 0$ as the tolerance for Picard iteration. The pressure head at time t^{j+1} is then given by $p^{j+1} = p^{j+1,k+1}$, with k the total number of Picard iterations to converge to ε_{PI}. The iterative scheme (6.24) is a general mixed-formulation Picard iteration, which results in perfect mass balance.

6.4.2   Cell-centered multigrid

Focussing on the k-th Picard iteration (6.24) at time t^{j+1}, the following elliptic PDE with variable coefficients is obtained, using simplified notation,

$$\begin{aligned}
\frac{\widetilde{C}}{\Delta t}\delta \widetilde{p} - \nabla \cdot \left(\widetilde{K} \nabla \delta \widetilde{p}\right) &= \widetilde{f} \quad \text{in} \quad \mathcal{D}, \\
\delta \widetilde{p} &= 0 \quad \text{in} \quad \Gamma_D \cup \Gamma_N,
\end{aligned} \tag{6.25}$$

with the known quantities

$$\widetilde{C} = C(p^{j+1,k}), \quad \widetilde{K} = K^{j+1,k} \quad \text{and} \quad \widetilde{f} = \nabla \cdot K^{j+1,k} \nabla p^{j+1,k} + \frac{\partial K}{\partial z}^{j+1,k} + f^{j+1} - \frac{\theta^{j+1,k} - \theta^j}{\Delta t},$$

and the unknown $\delta \widetilde{p} = \delta p^{j+1,k}$. To discretize the above problem, we use a cell-centered finite volume scheme for which the hydraulic conductivity at the cell-face is based on the harmonic averaging of the hydraulic conductivities from the adjacent cells, derived by the continuity of fluxes [68, 81].

For the discretization of (6.25), a uniform grid $\mathcal{D}_h$ on a unit square domain with the same mesh width $h = 1/M, M \in \mathbb{N}$ in both directions,

$$\mathcal{D}_h = \left\{(x_{i_1}, z_{i_2}); x_{i_1} = \left(i_1 - \frac{1}{2}\right)h, z_{i_2} = \left(i_2 - \frac{1}{2}\right)h, i_1, i_2 = 1, \dots, M\right\}, \tag{6.26}$$

is considered. For each *interior cell* (edges do not lie on a boundary) with center (x_{i_1}, z_{i_2}), denoted by $\mathcal{D}_h^{i_1,i_2}$, we obtain a five-point scheme,

$$c_{i_1,i_2}^h \delta\widetilde{p}_{i_1,i_2} + w_{i_1,i_2}^h \delta\widetilde{p}_{i_1-1,i_2} + e_{i_1,i_2}^h \delta\widetilde{p}_{i_1+1,i_2} + s_{i_1,i_2}^h \delta\widetilde{p}_{i_1,i_2-1} + n_{i_1,i_2}^h \delta\widetilde{p}_{i_1,i_2+1} = \widetilde{f}_{i_1,i_2}^h, \tag{6.27}$$

with

$$
\begin{aligned}
w_{i_1,i_2}^h &= -\frac{2}{h^2} \frac{\widetilde{K}_{i_1,i_2}\widetilde{K}_{i_1-1,i_2}}{\widetilde{K}_{i_1,i_2} + \widetilde{K}_{i_1-1,i_2}}, \quad
e_{i_1,i_2}^h = -\frac{2}{h^2} \frac{\widetilde{K}_{i_1,i_2}\widetilde{K}_{i_1+1,i_2}}{\widetilde{K}_{i_1,i_2} + \widetilde{K}_{i_1+1,i_2}}, \\
s_{i_1,i_2}^h &= -\frac{2}{h^2} \frac{\widetilde{K}_{i_1,i_2}\widetilde{K}_{i_1,i_2-1}}{\widetilde{K}_{i_1,i_2} + \widetilde{K}_{i_1,i_2-1}}, \quad
n_{i_1,i_2}^h = -\frac{2}{h^2} \frac{\widetilde{K}_{i_1,i_2}\widetilde{K}_{i_1,i_2+1}}{\widetilde{K}_{i_1,i_2} + \widetilde{K}_{i_1,i_2+1}}, \\
c_{i_1,i_2}^h &= -(w_{i_1,i_2}^h + e_{i_1,i_2}^h + n_{i_1,i_2}^h + s_{i_1,i_2}^h) + \frac{\widetilde{C}_{i_1,i_2}}{\Delta t},
\end{aligned}
$$

where, for instance, $\widetilde{K}_{i_1,i_2}$ is the diffusion coefficient associated with cell $\mathcal{D}_h^{i_1,i_2}$ and the source term $\widetilde{f}_{i_1,i_2}^h$ is an approximation of $\widetilde{f}$ in that cell. This scheme is modified appropriately for cells close to the boundary.

Next, we describe the multigrid method for solving the linear system arising from the above discretization. The multigrid hierarchy is based on uniform grid coarsening, i.e. the cell-width is doubled in each coarsening step in each direction. As the smoothing method, we use the lexicographic Gauss-Seidel iteration, and as the transfer operators between the fine and coarse grids a simple piece-wise constant prolongation operator, P_{2h}^h, is applied and its scaled adjoint is used as the restriction operator R_h^{2h} on the cell-centered grid. In classical stencil notation, these are written as,

$$
P_{2h}^h = \left]\begin{matrix} 1 & & 1 \\ & \star & \\ 1 & & 1 \end{matrix}\right[_{2h}^{h}, \qquad
R_h^{2h} = \frac{1}{4}\left[\begin{matrix} 1 & & 1 \\ & \star & \\ 1 & & 1 \end{matrix}\right]_{h}^{2h}, \tag{6.28}
$$

respectively, where $\star$ denotes the position of the cell center. The coarse grid operator is obtained via a direct discretization of the PDE operator on the coarse grid. For this discretization on a coarser grid, we need to appropriately define the diffusion coefficients on the coarse cell edges. The technique to define the suitable diffusion coefficient on a coarse cell edge is graphically described in Figure 4.1 and its caption. In [24], the $W(2,2)$-cycle was found to be a very robust and efficient multigrid cycling strategy, and, therefore, we also employ this cycle in our experiments. The number of multigrid iterations is based on the stopping criterion, $||\mathcal{L}_h \delta\widetilde{p}_h - \widetilde{f}_h||_\infty < \varepsilon_{MG}$, where $\mathcal{L}_h$ denotes the linear operator after the discretization of Equation (6.25) and $\varepsilon_{MG} > 0$.

We consider the modified Picard method in this work as it is widely adopted, although many modifications have been proposed to improve its robustness. For instance, the authors in [139] studied a spatio-temporal adaptive solution method to improve the numerical stability of the modified Picard iteration. Another interesting improvement was proposed in [140], where an Anderson acceleration was applied to improve the robustness and computational cost for the standard Picard iteration scheme. These improvements can easily be extended to the modified Picard-CCMG solver studied here. Also, there are a number of effective solution approaches based on Newton's method, see for e.g. [141–143]. These methods exhibit a quadratic convergence rate but are very sensitive to initial solution approximations.

6.4.3 Performance of the modified Picard-CCMG solver

We study the performance of the modified Picard-CCMG solver for a range of values of the parameters α, n and the effect of the heterogeneity of the hydraulic conductivity on the performance of the

solver. For this we consider an *infiltration problem* [143, 144] on a two-dimensional computational domain $\mathcal{D} = (0,1)^2$. The initial and boundary conditions are prescribed as follows:

$$p(x,z,0) = -0.4(1 - \exp(-80z)), \qquad p(x,1,t) = -0.4, \qquad (6.29)$$
$$p(x,0,t) = 0.1, \qquad \left.\frac{\partial p}{\partial x}\right|_{x=0,1} = 0.$$

The right-hand side is assumed to be zero, and we consider a final time $T_{final} = 0.1$ [h] (in hours). In Table 6.2, we provide a list of 20 values for α and n, used in the experiments. In total, we test 400 pairings of α and n. Parameters $\theta_s = 0.50$ and $\theta_r = 0.05$ are fixed, as they do not pose any problems for the convergence rate of the solver. The samples of hydraulic conductivity are generated according to (6.12), with $K_s^{(bl)} = 0.2$ [m/h] (in metres/hour) and the covariance is based on the two Matérn parameters from Table 6.1.

α					n				
0.2	0.4	0.6	0.8	1.0	1.1	1.2	1.3	1.4	1.5
1.2	1.4	1.6	1.8	2.0	1.6	1.7	1.8	1.9	2.0
2.2	2.4	2.6	2.8	3.0	2.2	2.4	2.6	2.8	3.0
3.2	3.4	3.6	3.8	4.0	3.2	3.4	3.6	3.8	4.0

Table 6.2: Set of α, n values used for benchmarking the modified Picard-CCMG solver.

A similar test was performed in [140] for a deterministic steady-state flow governed by the Richards' equation. We perform our experiments in a probabilistic framework. For a given pair α, n, we generate 64 random hydraulic conductivity fields and solve Richards' equation with conditions given in (6.29) for each sample. This is done as the number of multigrid iterations varies depending on the random realization of the hydraulic conductivity field. The cost of solving one instance of stochastic Richards' equation is expressed in terms of the total number of multigrid $W(2,2)-$cycles needed to solve the time-dependent problem. Here, by total number of $W(2,2)-$ cycles means the sum of multigrid iterations needed to reach T_{final}. For all experiments, we set the tolerances $\varepsilon_{PI}, \varepsilon_{MG} = 10^{-5}$. The solution method was terminated with failure when the maximum number of nonlinear iterations (set to 50) was exceeded at any time-step.

In figures 6.2-6.3, we show the average cost (average of 64 random realizations of the hydraulic conductivity) for solving the stochastic Richards' equation for four different combinations of spatial and temporal grid sizes and for two Matérn parameter sets, Φ_1 and Φ_2, listed in Table 6.1. The region in red in the figures denotes the (α, n) values for which the modified Picard-CCMG solver failed to converge at least once out of the 64 samples.

Based on numerical experiments, the performance of the modified Picard-CCMG solver for the stochastic Richards' equation can be summarized as follows:

- In general, the cost increases by decreasing n and increasing α. The cost of the solver rises steeply for $n < 1.5$ and $\alpha > 3.0$, and the cost increment with respect to the decrease in the value of n is more pronounced compared to the increase in α.

- While a spatio-temporal mesh refinement improves the robustness with respect to α and n, the improvement is less pronounced for n and may require a very fine mesh as $n \to 1$.

- For a given spatio-temporal mesh, the modified Picard-CCMG solver is less robust and more expensive for anisotropic hydraulic conductivity compared to the isotropic case. A similar (α, n)-robustness can be achieved for anisotropic cases by using a sufficiently refined mesh.

(a) $h = 1/32, \Delta t = 1/64$

(b) $h = 1/32, \Delta t = 1/128$

(c) $h = 1/64, \Delta t = 1/64$

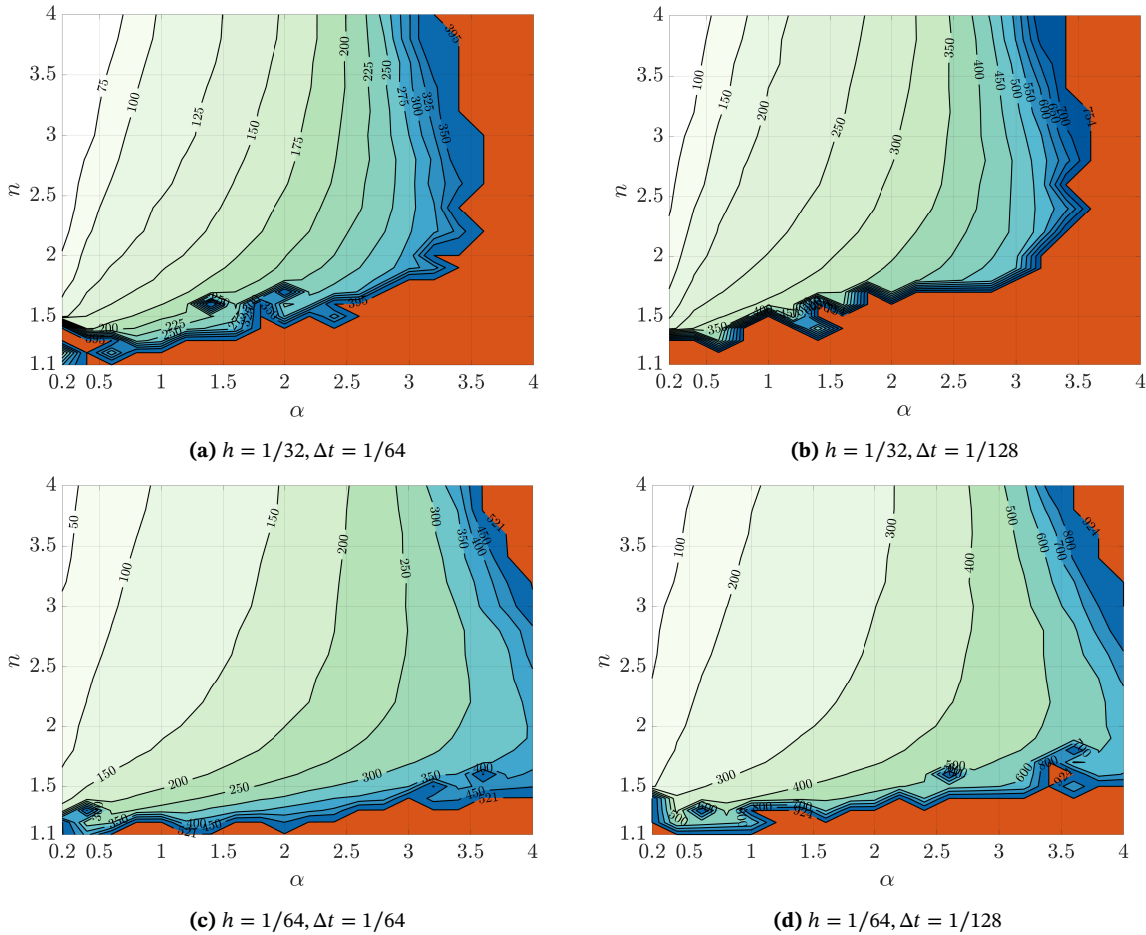**(d)** $h = 1/64, \Delta t = 1/128$

Figure 6.2: Contour plots of the average number of multigrid iterations needed to solve the infiltration problem using the modified Picard-CCMG solver for isotropic hydraulic conductivity fields generated using Φ_1.

The standard deviation contours for the cost show a similar behavior as the average cost contour and we observe a large standard deviation for the cost when $n < 1.5$ and $\alpha > 3.0$. In Figure 6.4, we present the number of samples (out of 64 samples), for which the solver did not converge for Φ_1 and Φ_2. For almost all samples convergence failed with α close to 4.0 and n close to 1.1.

A few remarks are in order. We point out that the (α, n)-cost map may vary depending on the type of boundary and initial conditions as well as on T_{final}. For instance, in the above experiments, an initially wet profile for the porous media was considered. We expect the performance of the modified Picard-CCMG solver to vary for problems in which infiltration takes place into an initially dry media and the convergence rates may depend on the values of θ_r and θ_s (see e.g. [145]). Furthermore, the robustness of the solver will also depend on the properties of the hydraulic conductivity field such as on the degree of heterogeneity and anisotropy.

6.5. Multilevel Monte Carlo with parametric continuation

We have observed in the preceding section that the total number of multigrid iterations increases rapidly with a decrease in the value of parameter n and an increase in α. We also noticed that the solver is less robust on a coarse spatio-temporal mesh. Therefore, when using the original MLMC estimator for a "difficult" (α, n) pair, a relatively fine spatio-temporal mesh will be required (and employed), even on the coarsest level of the MLMC hierarchy, resulting in an expensive estimator. To deal with this drawback, we propose an MLMC estimator based on the *parametric continuation*
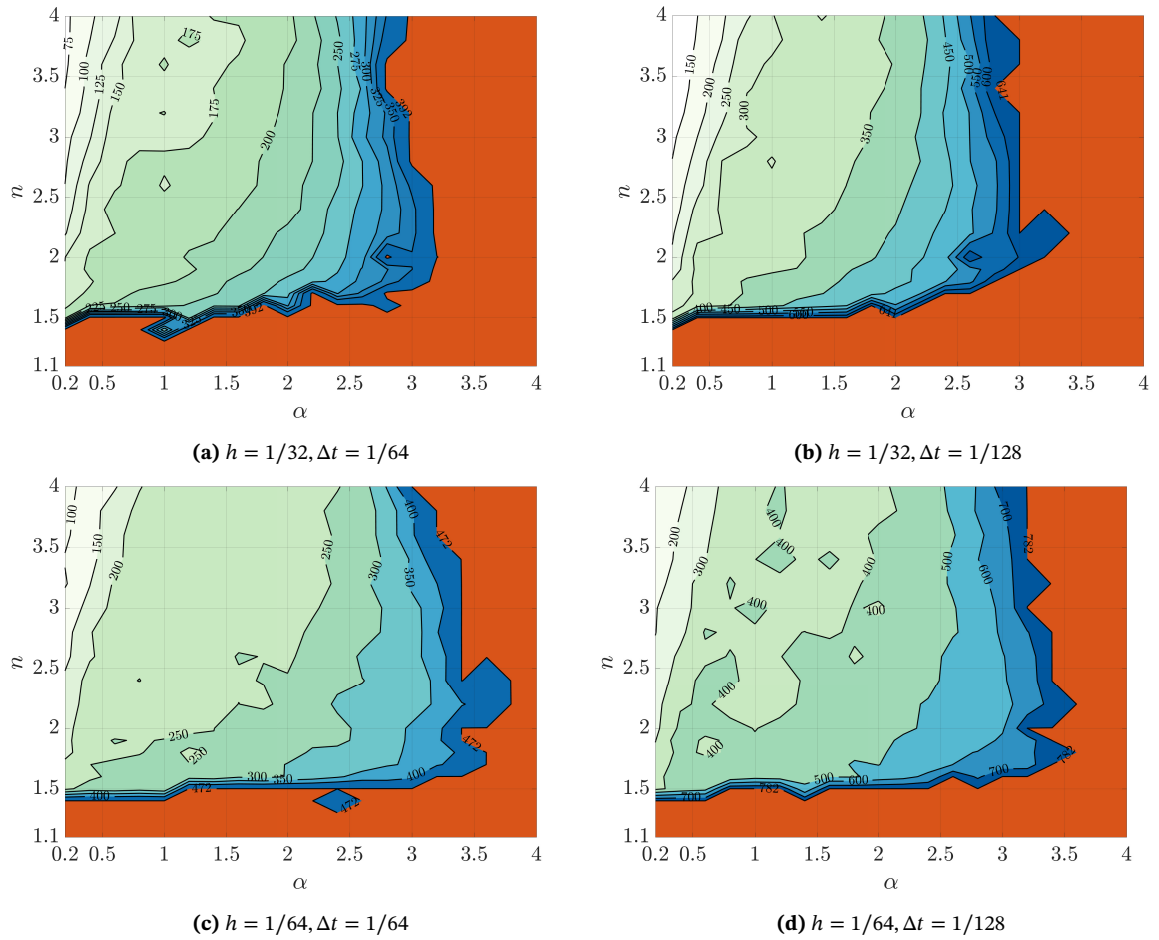
(a) $h = 1/32, \Delta t = 1/64$

(b) $h = 1/32, \Delta t = 1/128$

(c) $h = 1/64, \Delta t = 1/64$

(d) $h = 1/64, \Delta t = 1/128$

Figure 6.3: Contour plots of the average number of multigrid iterations needed to solve (6.29) using the modified Picard-CCMG solver for anisotropic hydraulic conductivity fields generated using Φ_2.

(a) $h = 1/64, \Delta t = 1/128, \Phi_1$

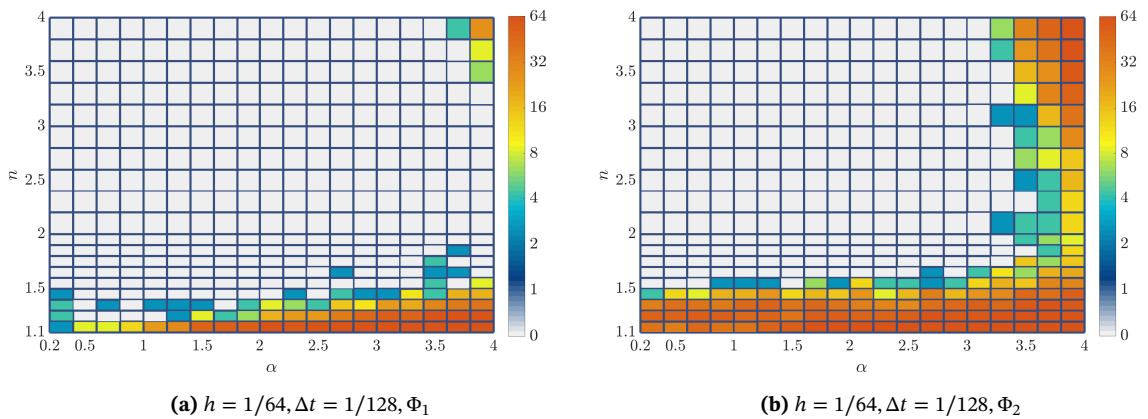**(b)** $h = 1/64, \Delta t = 1/128, \Phi_2$

Figure 6.4: Counting the number of samples (out of 64), for which the modified Picard-CCMG solver does not converge.

technique. In this approach, we solve the original problem only on the finest level of the MLMC hierarchy and *simplify* the parameter settings dictating the nonlinearities as we work on coarser levels. This allows us to include a comparatively coarser spatio-temporal mesh compared to the original MLMC estimator as simpler problems are solved on coarser levels.

This idea is motivated by continuation based multigrid solvers for nonlinear boundary value problems [128–130]. In the context of multigrid solvers, the continuation is commonly applied in the FMG-FAS (Full MultiGrid- Full Approximation Scheme) algorithm. In these algorithms, the continuation process is integrated with the FMG hierarchy, where the coarse grid solves the simplest problem and is used as a good first approximation for the next grid with a slightly more complicated problem. This process is repeated until the finest grid is reached where the target problem is solved. Although the continuation strategy works well for a large class of nonlinear problems, there is no guarantee that the simpler problem is close enough to the next difficult problem. One can use *bifurcation diagrams* to understand the solution dependence on nonlinearity dictating parameters. These diagrams can also reveal multiple branches and bifurcation points, where the solution differs greatly even if there is a slight perturbation in the parameter value. In such cases, an *arclength* procedure [131] can be applied to determine the appropriate perturbation size.

6.5.1 MLMC estimator

To explain the MLMC estimator, we consider the pressure head field at some final time T_{final} as the QoI. Further, we define a spatio-temporal hierarchy of grid levels $\{\mathcal{D}_\ell, \mathcal{T}_\ell\}_{\ell=0}^L$ using

$$h_\ell = \Delta t_\ell = \mathcal{O}(s^{-\ell} h_0), \tag{6.30}$$

where h_0 is the cell-width on the coarsest mesh $\mathcal{D}_0$ and $s > 0$ represents a grid refinement factor. We further define a hierarchy of parameter sets, $\{\boldsymbol{\Theta}_\ell\}_{\ell=0}^L$, where $\boldsymbol{\Theta}_L$ is the parameter set corresponding to the target (strongly nonlinear) problem to be solved. For instance, we can define a parametric hierarchy using the set of van Genuchten parameters, e.g. $\boldsymbol{\Theta}_\ell = \{\alpha_\ell^{(bl)}, n_\ell^{(bl)}\}$. The approximation of the pressure head on the level ℓ at T_{final} is denoted by $p_{h_\ell, \boldsymbol{\Theta}_\ell}$. Using the linearity of the expectation operator, one can define the expected value of the pressure head on the finest level, L, with the original parameter set, $\boldsymbol{\Theta}_L$, by the following telescopic sum:

$$\mathbb{E}[p_{h_L, \boldsymbol{\Theta}_L}] = \mathbb{E}[p_{h_0, \boldsymbol{\Theta}_0}] + \sum_{\ell=1}^L \mathbb{E}[p_{h_\ell, \boldsymbol{\Theta}_\ell} - p_{h_{\ell-1}, \boldsymbol{\Theta}_{\ell-1}}]. \tag{6.31}$$

Note that for $\boldsymbol{\Theta}_0 = \boldsymbol{\Theta}_1 = \ldots = \boldsymbol{\Theta}_L$, we have the standard MLMC estimator which solves the same problem on all levels. In terms of computational effort, it is cheaper to approximate $\mathbb{E}[p_{h_0, \boldsymbol{\Theta}_0}]$ by a standard Monte Carlo estimator, as the samples are computed on a coarse spatio-temporal mesh based on an "easy nonlinear parameter set $\boldsymbol{\Theta}_0$". Furthermore, the correction term, $\mathbb{E}[p_{h_\ell, \boldsymbol{\Theta}_\ell} - p_{h_{\ell-1}, \boldsymbol{\Theta}_{\ell-1}}]$, can be accurately determined using only a few samples as the level-dependent variance, $\mathcal{V}_\ell := \mathbb{V}[p_{h_\ell, \boldsymbol{\Theta}_\ell} - p_{h_{\ell-1}, \boldsymbol{\Theta}_{\ell-1}}]$, is typically small, since the random variables $p_{h_\ell, \boldsymbol{\Theta}_\ell}$ and $p_{h_{\ell-1}, \boldsymbol{\Theta}_{\ell-1}}$ are positively correlated. Note that the correlation will depend on the grid parameters h_ℓ and $h_{\ell-1}$ as well as on the difference between the nonlinear parameters $\boldsymbol{\Theta}_\ell$ and $\boldsymbol{\Theta}_{\ell-1}$. We will elaborate on this later on.

Each of the expectations in the MLMC estimator (6.31) can be independently computed using the standard MC simulation. We define a multilevel estimator, $\mathcal{E}_L^{ML}[p_{h_L, \boldsymbol{\Theta}_L}]$, constructed using a sum of $L + 1$ MC estimators:

$$\mathbb{E}[p_{h_L, \boldsymbol{\Theta}_L}] \approx \mathcal{E}_L^{ML}[p_{h_L, \boldsymbol{\Theta}_L}] := \sum_{\ell=0}^L \mathcal{E}_{N_\ell}^{MC}[p_{h_\ell, \boldsymbol{\Theta}_\ell} - p_{h_{\ell-1}, \boldsymbol{\Theta}_{\ell-1}}], \tag{6.32}$$

where $\mathcal{E}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell} - p_{h_{\ell-1},\Theta_{\ell-1}}]$ is the standard MC estimator obtained by averaging N_ℓ independent, identically distributed (i.i.d.) samples as

$$\mathcal{E}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell} - p_{h_{\ell-1},\Theta_{\ell-1}}] := \left(\frac{1}{N_\ell} \sum_{i=1}^{N_\ell} (p_{h_\ell,\Theta_\ell}(\omega_i) - p_{h_{\ell-1},\Theta_{\ell-1}}(\omega_i)) \right). \tag{6.33}$$

with ω_i denoting an event in the stochastic domain Ω and $p_{h_{-1},\Theta_{-1}} = 0$. It is expected that the number of MLMC samples $N_\ell \in \mathbb{N}$ forms a decreasing sequence for increasing ℓ. In order to keep the variance of the correction terms small, the MC samples, $p_{h_\ell,\Theta_\ell}(\omega_i) - p_{h_{\ell-1},\Theta_{\ell-1}}(\omega_i)$, should be based on the same random input ω_i for the simulation on two consecutive levels ℓ and $\ell - 1$.

Similarly, a multilevel estimator for the variance of the pressure head, $\mathbb{V}[p_{h_L,\Theta_L}]$, can be defined as

$$\mathbb{V}[p_{h_L,\Theta_L}] \approx \mathcal{V}_L^{ML}[p_{h_L,\Theta_L}] := \sum_{\ell=0}^{L} \mathcal{V}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell}] - \mathcal{V}_{N_\ell}^{MC}[p_{h_{\ell-1},\Theta_{\ell-1}}], \tag{6.34}$$

where the variance $\mathcal{V}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell}]$ is computed as

$$\mathcal{V}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell}] \approx \frac{1}{N_\ell - 1} \sum_{i=1}^{N_\ell} \left(p_{h_\ell,\Theta_\ell}(\omega_i) - \mathcal{E}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell}] \right)^2. \tag{6.35}$$

Again, the computational savings for the variance estimator (6.34) are obtained by computing individual variances $\mathcal{V}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell}]$ and $\mathcal{V}_{N_\ell}^{MC}[p_{h_{\ell-1},\Theta_{\ell-1}}]$ using the same random inputs $\{\omega_i\}_{i=1}^{N_\ell}$. The above variance estimator can be seen as an extension of the standard multilevel variance estimator proposed in [146].

For the multilevel estimators, an appropriate spatial interpolation procedure is required to combine expectations from all levels. Typically, the polynomial order of the interpolation scheme should be equal to or higher than the order of the discretization to avoid any additional dominant source of error. In some more detail, when using the estimator (6.32) to compute $\mathcal{E}_L^{ML}[p_{h_L,\Theta_L}]$, we begin by computing $\mathcal{E}_{N_0}^{MC}[p_{h_0,\Theta_0}]$ on the coarsest grid $\mathcal{D}_0$. This quantity is then interpolated to the next finer grid $\mathcal{D}_1$ and is added to the correction term $\mathcal{E}_{N_1}^{MC}[p_{h_1,\Theta_1} - p_{h_0,\Theta_0}]$ resulting in a two-level estimate $\mathcal{E}_1^{ML}[p_{h_1,\Theta_1}]$. This is again interpolated to the next grid level $\mathcal{D}_2$ and added to the next correction term $\mathcal{E}_{N_2}^{MC}[p_{h_2,\Theta_2} - p_{h_1,\Theta_1}]$. This process is repeated until the final level is reached.

6.5.2   Accuracy of MLMC estimator

Throughout this chapter, we use the L^2- based norm for the error analysis of the multilevel Monte Carlo estimator. We assume that the pressure head considered belongs to the functional space $L^2(\Omega, \mathcal{D})$ corresponding to the space of square-integrable measurable functions $p : \Omega \to L^2(\mathcal{D})$ for a previously defined probability space $(\Omega, \mathbb{F}, \mathbb{P})$. These spaces are equipped with the norm

$$\|p(\mathbf{x}, T, \omega)\|_{L^2(\Omega,\mathcal{D})} := \mathbb{E}\left[\|p(\mathbf{x}, T, \omega)\|_{L^2(\mathcal{D})}^2 \right]^{\frac{1}{2}} = \left(\int_\Omega \|p(\mathbf{x}, T, \omega)\|_{L^2(\mathcal{D})}^2 \, d\mathbb{P} \right)^{\frac{1}{2}}. \tag{6.36}$$

Similar to Section 2.3.1, the mean square error (MSE) in $\mathcal{E}_L^{ML}[p_{h_L,\Theta_L}]$ can then be expressed as the sum of the discretization and the sampling errors as

$$\left\| \mathbb{E}[p_{\Theta_L}] - \mathcal{E}_L^{ML}[p_{h_L,\Theta_L}] \right\|_{L^2(\Omega,\mathcal{D})}^2 \leq \left\| \mathbb{E}[p_{\Theta_L}] - \mathbb{E}[p_{h_L,\Theta_L}] \right\|_{L^2(\mathcal{D})}^2 + \left\| \mathbb{E}[p_{h_L,\Theta_L}] - \mathcal{E}_L^{ML}[p_{h_L,\Theta_L}] \right\|_{L^2(\Omega,\mathcal{D})}^2. \tag{6.37}$$

Both errors in the MLMC estimator can be dealt with separately. The discretization error can be quantified as:

$$\left\|\mathbb{E}[p_{\boldsymbol{\Theta}_L}] - \mathbb{E}[p_{h_L,\boldsymbol{\Theta}_L}]\right\|_{L^2(\mathcal{D})} \le c_0 h_L^a, \qquad a > 0, \tag{6.38}$$

where c_0 is a constant independent of h_L but depending on the parameter set $\boldsymbol{\Theta}_L$. The rate a typically depends on the regularity of the PDE and the accuracy of the discretization. The next task is to bound the sampling errors. As the MLMC estimator $\mathcal{E}_L^{ML}[p_{h_L,\boldsymbol{\Theta}_L}]$ is composed of $L+1$ independent MC estimators, the sampling error in the MLMC estimator is just the sum of sampling errors from the individual MC estimators. Therefore,

$$\left\|\mathbb{E}[p_{h_L,\boldsymbol{\Theta}_L}] - \mathcal{E}_L^{ML}[p_{h_L,\boldsymbol{\Theta}_L}]\right\|_{L^2(\Omega,\mathcal{D})}^2 = \sum_{\ell=0}^{L} \frac{\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}}{N_\ell}, \tag{6.39}$$

see [30, 147] for a proof. Obtaining a bound on the level-variance $\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}$ is more involved due to its dependence on the grid size h_ℓ as well as on the nonlinearity parameter set $\boldsymbol{\Theta}_\ell$. We numerically estimate it by

$$
\begin{aligned}
\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})} &= \left\|\mathbb{V}[p_{h_\ell,\boldsymbol{\Theta}_\ell} - p_{h_{\ell-1},\boldsymbol{\Theta}_{\ell-1}}]\right\|_{L^2(\mathcal{D})} \\
&\approx \frac{1}{N_\ell - 1} \sum_{i=1}^{N_\ell} \int_{\mathcal{D}} \left(\mathcal{E}_{N_\ell}^{MC}[p_{h_\ell,\boldsymbol{\Theta}_\ell} - p_{h_{\ell-1},\boldsymbol{\Theta}_{\ell-1}}] - (p_{h_\ell,\boldsymbol{\Theta}_\ell}(\omega_i) - p_{h_{\ell-1},\boldsymbol{\Theta}_{\ell-1}}(\omega_i))\right)^2
\end{aligned} \tag{6.40}
$$

Recall from Section 2.3.1, to achieve a tolerance of ε, one needs to ensure that

$$\left\|\mathbb{E}[p_{\boldsymbol{\Theta}_L}] - \mathcal{E}_L^{ML}[p_{h_L,\boldsymbol{\Theta}_L}]\right\|_{L^2(\Omega,\mathcal{D})}^2 \le (c_0 h_L^a)^2 + \sum_{\ell=0}^{L} \frac{\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}}{N_\ell} < \varepsilon^2. \tag{6.41}$$

The total cost of the MLMC estimator can be expressed as $\mathcal{W}_L^{ML} = \sum_{\ell=0}^{L} N_\ell \mathcal{W}_\ell$, where $\mathcal{W}_\ell = \mathcal{O}(h_\ell^{-\gamma})$ corresponds to the cost of computing one sample on level ℓ. For time-dependent problems, the rate $\gamma \ge d+1$, with d the number of spatial dimensions. As proposed in [10, 28], the number of samples at different levels is typically derived by minimizing the total cost such that the sampling error of the MLMC estimator reduces below ε^2, i.e.,

$$\min\left(\sum_{\ell=0}^{L} N_\ell \mathcal{W}_\ell\right) \quad \text{s.t} \quad \sum_{\ell=0}^{L} \frac{\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}}{N_\ell} = \varepsilon^2. \tag{6.42}$$

Using the standard Lagrange multiplier approach [10], gives us

$$N_\ell = \varepsilon^{-2}\left(\sum_{k=0}^{L} \sqrt{\|\mathcal{V}_k\|_{L^2(\mathcal{D})} \mathcal{W}_k}\right)\sqrt{\frac{\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}}{\mathcal{W}_\ell}}, \tag{6.43}$$

and hence the total cost to obtain a tolerance of ε is given by

$$\mathcal{W}_L^{ML}(\varepsilon) = \sum_{\ell=0}^{L} N_\ell \mathcal{W}_\ell = \varepsilon^{-2}\left(\sum_{\ell=0}^{L} \sqrt{\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})} \mathcal{W}_\ell}\right)^2. \tag{6.44}$$

In the above formula, the product $\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})} \mathcal{W}_\ell$ determines the cost contribution from any level ℓ. For instance, if the product decays with increasing ℓ, the dominant cost comes from the coarsest level whereas if the product grows with ℓ, the dominant contribution comes from the finest level.

Remark 6.5.1 *The optimal number of samples given in (6.43) is based on a pre-defined hierarchy of parameters $\{\Theta_\ell\}_{\ell=0}^L$. A more general approach is to find N_ℓ along with the parameter set $\{\Theta_\ell\}_{\ell=0}^L$ for which the total cost of the MLMC estimator is minimum. Solving such optimization problem analytically is non-trivial. Furthermore, numerically obtaining the best values for Θ_ℓ can also be highly expensive. In the numerical experiments section, we will discuss some heuristics that can be applied to find Θ_ℓ. We further point out that the multi-index Monte Carlo (MIMC) estimator [34] is also highly suited for this problem. For MIMC estimator, a two-dimensional hierarchy of indices based on h_ℓ and Θ_ℓ can be constructed to achieve more effective variance reduction.*

MLMC algorithm with parametric continuation

To compute the estimator $\mathcal{E}_L^{ML}[p_{h_L,\Theta_L}]$, the standard MLMC algorithm from [10, 28] cannot be directly employed as it requires solving the same problem on all grid levels. Here, we describe a *modified version* of the standard MLMC technique to compute $\mathcal{E}_L^{ML}[p_{h_L,\Theta_L}]$. This algorithm assumes that the total number of levels in the MLMC hierarchy and the values of the nonlinearity parameters Θ_ℓ for all levels are known in advance. The algorithm can be described by the following steps:

Algorithm 3 *PC_MLMC* algorithm

1: Fix the tolerance ε, $\mathcal{D}_\ell$, $\mathcal{T}_\ell$, Θ_ℓ and warm-up samples N_ℓ^* for $\ell = 0, 1, 2, ..., L$.
2: Compute quantities $\mathcal{E}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell} - p_{h_{\ell-1},\Theta_{\ell-1}}]$ and $\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}$ using samples $N_\ell = N_\ell^*$ for all levels.
3: Update N_ℓ using the formula (6.43) for all levels.
4: Compute additional samples and update $\mathcal{E}_{N_\ell}^{MC}[p_{h_\ell,\Theta_\ell} - p_{h_{\ell-1},\Theta_{\ell-1}}]$ and $\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}$ for all levels.
5: Perform steps 3-4 until no additional samples are needed on any level.

In the above algorithm, the value of N_ℓ^* should not be set too high, especially not on the finest level, in order to avoid oversampling. Further, the cost per sample $\mathcal{W}_\ell$ can also be estimated "on-the-fly" by averaging the CPU times from the computation of warm-up samples.

6.6. Numerical experiments

We evaluate the performance of the new MLMC estimator and study the improvements with respect to the standard MLMC estimator. For all the experiments, we use the infiltration problem with conditions given in (6.29), however, with $T_{final} = 0.2$ [h] and the two Matérn covariance parameters from Table 6.1. We employ a geometric hierarchy of spatio-temporal grids with refinement factor $s = 2$ in (6.30) and we use $h_\ell = \Delta t_\ell$. For all experiments, the following baseline values are prescribed, $K_s^{(bl)} = 0.2$ [m/h], $\theta_s^{(bl)} = 0.5$ and $\theta_r^{(bl)} = 0.05$; different baseline values for $\alpha^{(bl)}$ and $n^{(bl)}$ are studied. The uncertainty in the soil parameters is defined according to the values presented in Table 6.3. The sampling and upscaling procedure for a Gaussian random field is already described in Chapter 3. The sampling of random fields with uniform marginal is described in Section 6.3.

Quantity	Uncertainty
$Z(\mathbf{x},\omega)$	$\mathcal{N}(0, C_\Phi)$
$\varepsilon_{\theta_s}(\mathbf{x},\omega)$	$\mathcal{U}(-0.05, 0.05, C_\Phi)$
$\varepsilon_{\theta_r}(\mathbf{x},\omega)$	$\mathcal{U}(-0.005, 0.005, C_\Phi)$
$\varepsilon_\alpha(\mathbf{x},\omega)$	$\mathcal{U}(-0.2, 0.2, C_\Phi)$
$\varepsilon_n(\mathbf{x},\omega)$	$\mathcal{U}(-0.05, 0.05, C_\Phi)$

Table 6.3: Description of uncertainty for different soil parameters.

Note that the above stochastic model is extremely high-dimensional as it comprises five independent random fields. For each random field the degree of freedom is equal to the number of grid points in the sampling mesh. The dimensionality can be reduced using the KL expansion method, however as we use random fields with small correlation lengths, we will still need to use a very large number of KL modes for an accurate representation of these random fields.

6.6.1 Convergence of discretization bias

We begin by analyzing the reduction of the discretization error $||p_{h_\ell} - p_{h_{\ell-1}}||_{L^2(\Omega,\mathcal{D})}$ with respect to mesh refinement for different baseline values of $\alpha^{(bl)}$ and $n^{(bl)}$. The relative error is used to bound the exact discretization bias as

$$||p - p_{h_\ell}||_{L^2(\Omega,\mathcal{D})} \leq \frac{||p_{h_\ell} - p_{h_{\ell-1}}||_{L^2(\Omega,\mathcal{D})}}{s^a - 1}, \tag{6.45}$$

where a is the convergence rate defined in (6.38). The relative errors for, Φ_1 and Φ_2 are presented in the left and right pictures in Figure 6.5, respectively. For both cases a convergence rate close to first-order is observed, i.e. $a \approx 1$. The convergence rate typically depends on the order of the spatio-temporal discretization scheme as well as on the smoothness parameter ν_c in the covariance function. In fact, the dominant error comes from the first-order accurate backward Euler time discretization. The magnitude of the error grows with increasing $\alpha^{(bl)}$ and reduces with increasing $n^{(bl)}$ values. Note that for the most difficult cases, $n^{(bl)} = 1.45, \alpha^{(bl)} = 3.0$ for Φ_1 and $n^{(bl)} = 1.55, \alpha^{(bl)} = 2.8$ for Φ_2, the convergent solutions are obtained from $h_\ell = 1/64$ onwards.
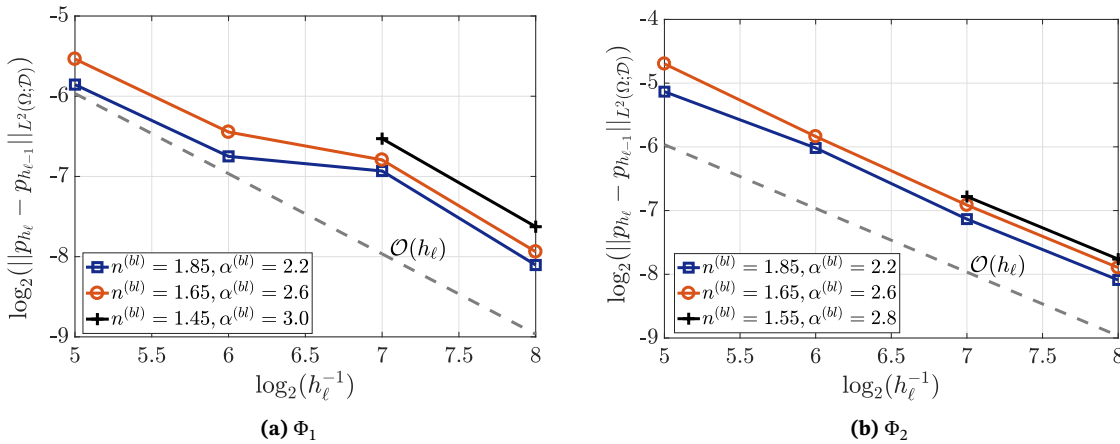

Figure 6.5: Convergence of discretization bias with mesh refinement for different baseline values of $\alpha^{(bl)}$ and $n^{(bl)}$.

6.6.2 MLMC simulation

Here, we describe the algorithm to compute the multilevel estimator $\mathcal{E}_L^{ML}[p_{h_L,\Theta_L}]$. We perform the MLMC simulations for two test cases based on Φ_1 and Φ_2, respectively. The original problem for Φ_1 uses $\Theta_L = \{\alpha_L^{(bl)}, n_L^{(bl)}\} = \{3.0, 1.45\}$ and for Φ_2 the original problem is based on $\Theta_L = \{2.8, 1.55\}$.

We first investigate the correlations for the pressure head profiles when the baseline values for $\alpha^{(bl)}$ and $n^{(bl)}$ are varied however employing the same random fields. In Figure 6.6, we compare three pressure head solutions with different baseline values, and with the same random fields, $Z(\mathbf{x}, \omega), \varepsilon_{\theta_s}(\mathbf{x}, \omega), \varepsilon_{\theta_r}(\mathbf{x}, \omega), \varepsilon_\alpha(\mathbf{x}, \omega)$ and $\varepsilon_n(\mathbf{x}, \omega)$ (see Section 6.3). Clearly, the pressure head profile becomes more *diffusive* when "easier" parameters are prescribed. We also compare the cross sections of the pressure head profiles at $x = 0.5$ in Figure 6.7. For reference, we use the solution on the fine grid $h = \Delta t = 1/128$ (black solid line) and compare it with different pairs of $n^{(bl)}$ and $\alpha^{(bl)}$
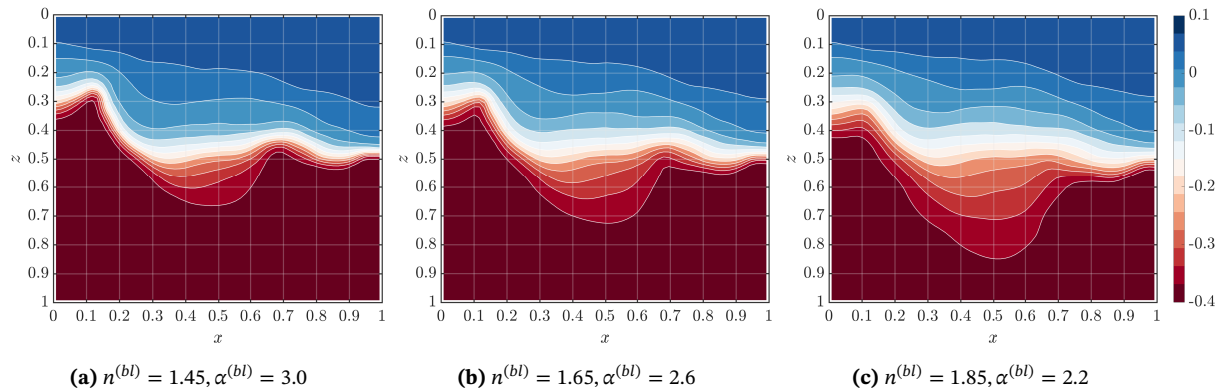
(a) $n^{(bl)} = 1.45, \alpha^{(bl)} = 3.0$ **(b)** $n^{(bl)} = 1.65, \alpha^{(bl)} = 2.6$ **(c)** $n^{(bl)} = 1.85, \alpha^{(bl)} = 2.2$

Figure 6.6: Comparison of pressure head fields at $T_{final} = 0.2$ [h] for different baselines values of the parameters, $(n^{(bl)}, \alpha^{(bl)})$ but with the same random fields $Z, \varepsilon_{\theta_s}, \varepsilon_{\theta_r}, \varepsilon_\alpha, \varepsilon_n$. Solutions are based on $h = \Delta t = 1/64$ and the Matérn parameter set Φ_1.
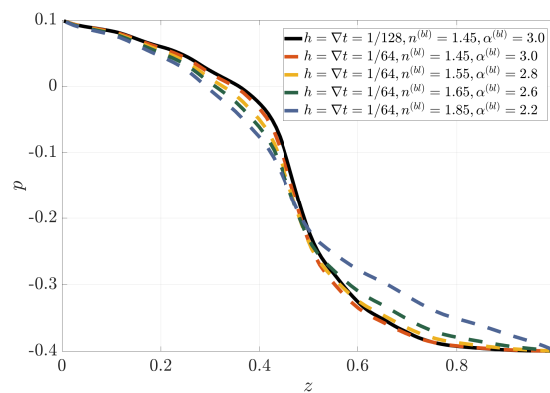

Figure 6.7: Comparison of cross sections of the pressure heads from Figure 6.6 at $x = 0.5$.

values on the next coarse grid $h = \Delta t = 1/64$. The profiles with the same $(n^{(bl)}, \alpha^{(bl)})$-values are very close and the deviation increases as the two parameters are set to "easier" values. Thus, we can conclude that the correlation decays as the difference between the baseline values of the nonlinear parameters widens.

Next, we study the behavior of the level-dependent variance $\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}$ when using the parametric continuation approach. For this we define the so-called *parametric continuation variables*, $\nabla\alpha = \alpha_\ell - \alpha_{\ell-1}$ and $\nabla n = n_{\ell-1} - n_\ell$, with the purpose to reduce the nonlinearity when processing coarse grids. In Figure 6.8, we plot $\|\mathcal{V}_\ell\|_{L^2(\mathcal{D})}$ computed using (6.40) for different $(\nabla\alpha, \nabla n)$ pairs for the two Matérn parameter sets. The original problem is solved with $h_L = 1/256$. The parameter sets Θ_L and Θ_ℓ, for $\ell = L - 1, L - 2, ..., 0$, are obtained by employing $\nabla\alpha$ and ∇n. The black line represents the variance when the same problem is solved on all levels, i.e. $\nabla\alpha = \nabla n = 0$, corresponding to the original MLMC estimator. Using the original approach, we can only process three levels in the MLMC hierarchy. The red and blue lines in the figure correspond to the variance which is computed using $\nabla\alpha = 0.05, \nabla n = 0.1$ and $\nabla\alpha = 0.1, \nabla n = 0.2$, respectively. For these two cases, we can incorporate a larger number of coarse levels, up to $h_0 = 1/16$, as milder nonlinear problems are solved on these coarse levels. Furthermore, for levels $\ell < L$ the variance is smaller, compared to the case without continuation (where $\nabla\alpha = \nabla n = 0$) which will result in a lower number of samples on these levels. Here, we wish to highlight the fact that choosing optimal values for $\nabla\alpha$ and ∇n is important. For example, when $\alpha = 0.1, \nabla n = 0.2$, the variance on the finest level increases as compared to the variance found with the original MLMC approach. Due to this, an
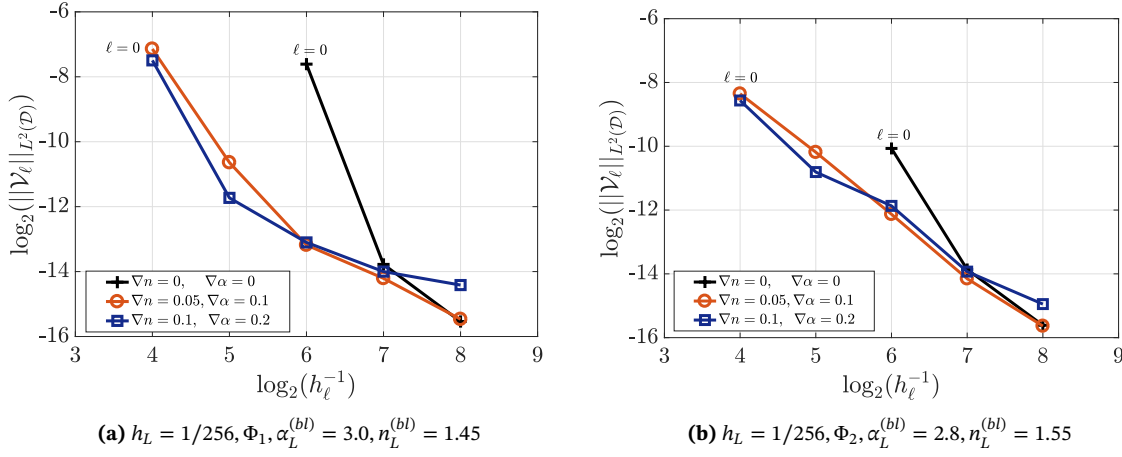
(a) $h_L = 1/256, \Phi_1, \alpha_L^{(bl)} = 3.0, n_L^{(bl)} = 1.45$

(b) $h_L = 1/256, \Phi_2, \alpha_L^{(bl)} = 2.8, n_L^{(bl)} = 1.55$

Figure 6.8: Comparison of level-dependent variances using different pairs of $\nabla\alpha, \nabla n$ for the two Matérn parameters.

increasing number of samples will be needed on the finest level, compared to the original MLMC estimator, which is undesirable as it may result in an expensive estimator. On the other hand, for a smaller perturbation $\nabla\alpha = 0.05, \nabla n = 0.1$, the magnitude of the variance on the finest level is similar to that of the original MLMC estimator. Therefore, the number of samples on the finest level will be more-or-less similar to the original MLMC estimator.

Alternatively, one can avoid high variance on the finest level by using zero perturbations on the two finest levels, i.e. $\nabla\alpha_L = \nabla\alpha_{L-1} = 0$ and $\nabla n_L = \nabla n_{L-1} = 0$ and choosing non-zero perturbations on next coarser levels from $L-2$ onwards. This way, we solve the original problem on the two finest levels.

Now we apply the parametric continuation based MLMC estimator denoted by PC_MLMC, to compute the mean and variance of the pressure head field and also perform comparisons with respect to the standard MLMC estimator which is denoted by Std_MLMC. For this, we use the two previously discussed test cases: isotropic covariance Φ_1 with baseline values $n^{(bl)} = 1.45, \alpha^{(bl)} = 3.0$ and anisotropic covariance Φ_2 with baseline values $n^{(bl)} = 1.55, \alpha^{(bl)} = 2.8$. For simplicity, we use the continuation variables $\nabla\alpha = 0.05, \nabla n = 0.1$ for both the isotropic and anisotropic cases. We compare the number of samples needed on different grids for three values of the tolerances. For the Std_MLMC estimator, the coarsest possible level is $h_0 = \Delta t_0 = 1/64$, whereas for the PC_MLMC, we use $h_0 = \Delta t_0 = 1/16$. As the discretization error shows a first-order decay (see Figure 6.5), we set the tolerance $\varepsilon = \mathcal{O}(h_L)$. We use Algorithm 3 to reduce the sampling error to ε. In Table 6.4, the two estimators for the isotropic Matérn parameter Φ_1 are compared. Due to the sample optimization strategy (6.43), a large number of samples is shifted to coarser grids when using the PC_MLMC estimator. Furthermore, a fewer number of samples are required for the PC_MLMC estimator compared to the Std_MLMC, even on the finest level. This is due the fact that the sum $\sum_{\ell=0}^{L} \sqrt{||\mathcal{V}_\ell||_{L^2(\mathcal{D})} \mathcal{W}_\ell}$ for the PC_MLMC estimator is slightly smaller than for Std_MLMC. Moreover, a large computational gain is induced by the reduction in the number of samples on grid $h_\ell = 1/64$, for instance, for $\varepsilon = 0.005$, the number of samples reduced from 438 to 35 when using the parametric continuation. In Figure 6.9 (a) the CPU times for the two estimators are also compared. We observe a speed-up of about a factor of three for $\varepsilon = 0.005$.

A similar test is performed for the anisotropic problem. The number of samples for different tolerances are provided in Table 6.5 and the CPU times in Figure 6.9 (b). Again some improvement in computation times is observed, although the gain is not as pronounced as for the first problem. This is due to the fact that the second case uses simpler baseline values $n^{(bl)} = 1.55, \alpha^{(bl)} = 2.8$ and the cost reduction with parameter simplification is not very rapid. For the isotropic case
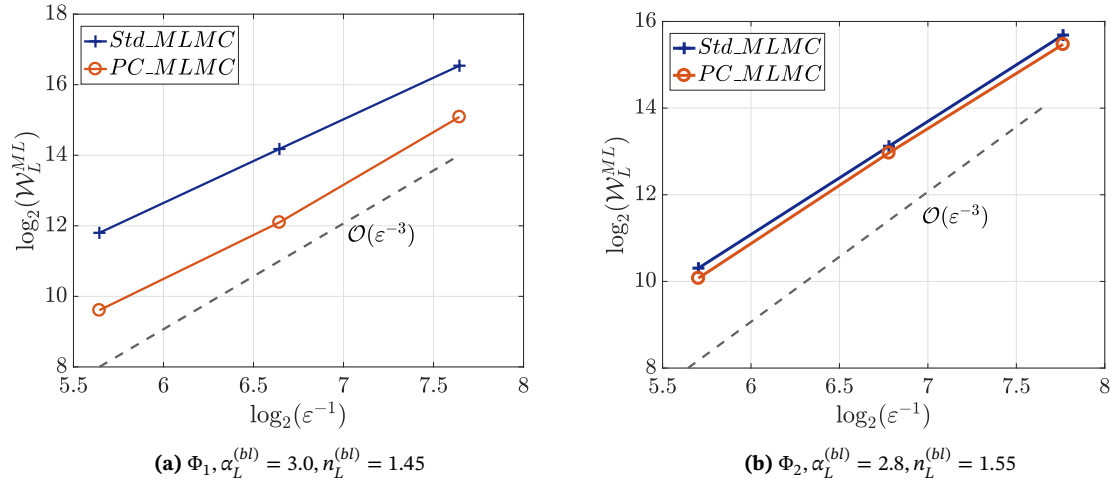
(a) $\Phi_1, \alpha_L^{(bl)} = 3.0, n_L^{(bl)} = 1.45$ **(b)** $\Phi_2, \alpha_L^{(bl)} = 2.8, n_L^{(bl)} = 1.55$

Figure 6.9: Comparison of CPU times $\mathcal{W}_L^{ML}$ (sec) for two different estimators.

with $n^{(bl)} = 1.45, \alpha^{(bl)} = 3.0$, the cost decay is more rapid with parameter simplification. This is more evident from the cost map in Figure 6.3, where we see more dense contour lines around $n^{(bl)} = 1.45$. Therefore, the parametric continuation approach is very effective when a strongly nonlinear stochastic problem needs to be solved.

We also wish to highlight the fact that both MLMC estimators are *optimal* since the cost scales as $\mathcal{O}(\varepsilon^{-3})$, which is similar to the computational complexity of solving one deterministic problem on the finest grid, i.e. $\mathcal{O}(h_L^{-3})$ and $h_L = \mathcal{O}(\varepsilon)$.

h_ℓ	$N_\ell(h_L = 1/64, \varepsilon = 0.02)$		$N_\ell(h_L = 1/128, \varepsilon = 0.01)$		$N_\ell(h_L = 1/256, \varepsilon = 0.005)$	
	Std_MLMC	*PC_MLMC*	*Std_MLMC*	*PC_MLMC*	*Std_MLMC*	*PC_MLMC*
1/16	−	115	−	459	−	1833
1/32	−	11	−	44	−	176
1/64	28	3	110	9	438	35
1/128	−	−	5	2	18	8
1/256	−	−	−	−	4	3

Table 6.4: Comparison of number of samples needed to achieve tolerances ε using the standard MLMC (*Std_MLMC*) and parametric continuation MLMC (*PC_MLMC*) estimators for $\Phi_1, n_L^{(bl)} = 1.45, \alpha_L^{(bl)} = 3.0$. Entries with symbol (−) indicate zero samples needed for that grid.

h_ℓ	$N_\ell(h_L = 1/64, \varepsilon = 0.0184)$		$N_\ell(h_L = 1/128, \varepsilon = 0.0092)$		$N_\ell(h_L = 1/256, \varepsilon = 0.0046)$	
	Std_MLMC	*PC_MLMC*	*Std_MLMC*	*PC_MLMC*	*Std_MLMC*	*PC_MLMC*
1/16	−	96	−	427	−	1659
1/32	−	18	−	77	−	301
1/64	9	4	36	16	140	60
1/128	−	−	4	4	15	12
1/256	−	−	−	−	3	3

Table 6.5: Comparison of number of samples needed to achieve tolerances ε using the standard MLMC (*Std_MLMC*) and parametric continuation MLMC (*PC_MLMC*) estimators for $\Phi_2, n^{(bl)} = 1.55, \alpha^{(bl)} = 2.8$.

In the last part of this section, we validate the stochastic moments computed using the proposed estimator. It is expected that the mean pressure field computed using the two MLMC estimators

should converge to a similar solution for a given tolerance. In Figure 6.10, the mean pressure head profile for the isotropic case is shown. It is computed using the number of samples from Table 6.4, with $\varepsilon = 0.005$. For a closer inspection, we also compare the mean pressure head profiles at $x = 0.5$. Similarly, the mean profile for the anisotropic case is presented in Figure 6.11, using the number of samples from Table 6.5 with $\varepsilon = 0.0046$. We see good agreement between the mean profiles computed from the two estimators. The isotropic case exhibits a seemingly smoother transition from the saturated to the unsaturated zone compared to the anisotropic problem. In Figure 6.12 we also present the variance field for the isotropic test case, computed using the multilevel variance estimator $\mathcal{V}_L^{ML}[p_{h_L,\boldsymbol{\Theta}_L}]$ given in (6.34). The two variance fields are very similar, although some discrepancy in the magnitude is observed. This is due to the fact that the two variance fields are computed using the samples based on the error analysis of $\mathcal{E}_L^{ML}[p_{h_L,\boldsymbol{\Theta}_L}]$ (from Table 6.4) and not on the error analysis of $\mathcal{V}_L^{ML}[p_{h_L,\boldsymbol{\Theta}_L}]$. Thus, the two variance estimates may have different tolerances resulting in this slight mismatch. Readers are referred to [146] for detailed error analysis of the multilevel variance estimator.

The results from the two estimators also showed good agreement with the plain Monte Carlo solutions performed on the grid $h_L = 1/128$. This is done in order to verify that a proper upscaling of the random fields on coarser levels is carried out while using the MLMC estimator. These results are omitted for the sake of brevity.
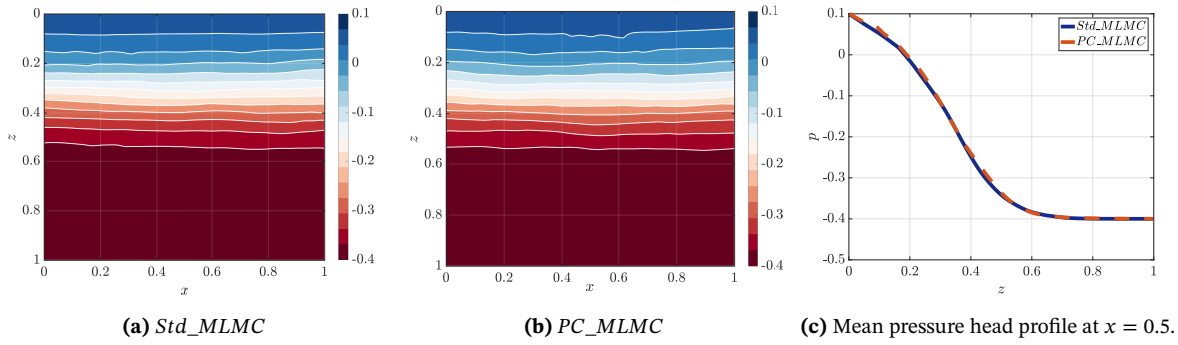

(a) *Std_MLMC*   **(b)** *PC_MLMC*   **(c)** Mean pressure head profile at $x = 0.5$.

Figure 6.10: Comparison of mean pressure head field for $\Phi_1, \alpha_L^{(bl)} = 3.0, n_L^{(bl)} = 1.45, T_{final} = 0.2$ [h] computed using the two MLMC estimators with finest level $h_L = \Delta t_L = 1/256$ and $\varepsilon = 0.005$.
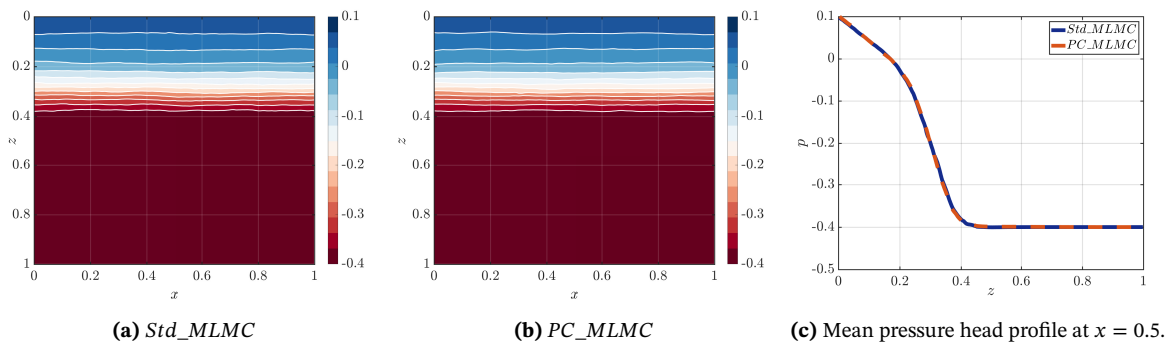

(a) *Std_MLMC*   **(b)** *PC_MLMC*   **(c)** Mean pressure head profile at $x = 0.5$.

Figure 6.11: Comparison of mean pressure head field for $\Phi_2, \alpha_L^{(bl)} = 2.8, n_L^{(bl)} = 1.55, T_{final} = 0.2$ [h] computed using the two MLMC estimators with finest level $h_L = \Delta t_L = 1/256$ and $\varepsilon = 0.0046$.

6.7. Conclusions

In this work, an efficient uncertainty propagation method for a high-dimensional stochastic extension of Richards' equation was proposed. All the soil parameters were treated as unknown and
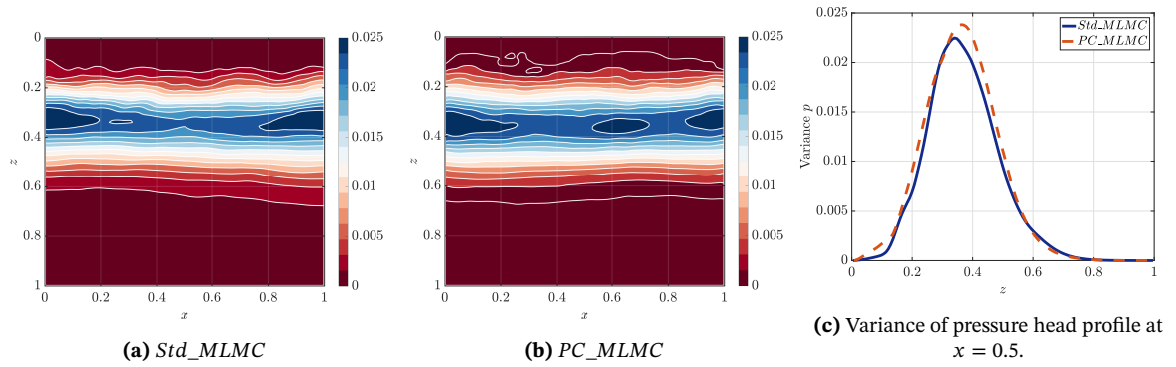
(a) Std_MLMC

(b) PC_MLMC

(c) Variance of pressure head profile at $x = 0.5$.

Figure 6.12: Comparison of the variance of the pressure head field for $\Phi_1, \alpha_L^{(bl)} = 3.0, n_L^{(bl)} = 1.45, T_{final} = 0.2$ [h] computed using the two MLMC variance estimators with finest level $h_L = \Delta t_L = 1/256$.

modeled as random fields with appropriate marginal distributions. We also studied a modified Picard iteration and cell-centered multigrid method for solving the nonlinear systems with heterogeneous coefficients. We found that the combined solver is robust for a wide parameter range and the performance further improves with spatio-temporal refinements. This combination of solvers is general, therefore, its robustness can be further improved by incorporating adaptive time stepping or by combining with other advanced techniques, for instance, using the Anderson acceleration proposed in [140].

For computing the statistical moments of the solution of Richards' equation, a parametric continuation technique based multilevel Monte Carlo estimator was proposed. This estimator is very practical for this problem, as it requires solving the strongly nonlinear problem only on the finest level, where the solver is robust and uses simpler nonlinear problems on the coarse grid levels for a variance reduction. For the stochastic Richards' equation, the proposed estimator is more prominent regarding the computational gains compared to the standard MLMC method if the problem is strongly nonlinear. In general, this estimator is also applicable to other parameter dependent nonlinear PDEs. One of the research problems that need to be addressed is finding a computationally viable way of obtaining optimal step sizes for the nonlinear parameters used in continuation.

7

Stochastic turbulence modeling

This chapter[1] presents a multilevel Monte Carlo method for quantifying model-form uncertainties associated with the Reynolds-Averaged Navier-Stokes (RANS) simulations. Two, high-dimensional, stochastic extensions of the RANS equations are considered to demonstrate the applicability of the MLMC method. The first approach is based on global perturbation of the baseline eddy viscosity field using a lognormal random field. A more general second extension is considered where the entire Reynolds Stress Tensor is perturbed while maintaining realizability. The MLMC experiments are performed for two fundamental flows along with comparisons in terms of cost and accuracy with plain Monte Carlo. We also discuss possibilities when an optimal multilevel estimator can be attained for which the cost scales with the same order as a single CFD solve on the finest grid level.

7.1. Introduction

The Reynolds-Averaged Navier-Stokes (RANS) equations combined with turbulence closure models are widely utilized in engineering to predict flows with high Reynolds numbers. These turbulence closure models are used to obtain an approximate Reynolds stress tensor that is responsible for coupling the mean flow with turbulence. Although many turbulence models exist in the literature, there is no single model that generalizes well to all classes of turbulent flows [148, 149]. Specifically, the performance depends on the modeling assumptions and the type of flow used to calibrate the so-called closure coefficients that are needed as inputs to a turbulence model.

Since the dominant source of error in the flow prediction comes from the turbulence modeling, a number of approaches have already been developed for the model-form uncertainty quantification (UQ) of RANS simulations, see e.g. [150, 151] for recent reviews. The majority of these approaches is based on the perturbation of baseline RANS models. One way to achieve this is by injecting uncertainties in the closure coefficients [22, 152–154] of the turbulence models. Another more general physics-based approaches exist, which typically introduces randomness directly into the modeled Reynolds Stress Tensor (RST), either by perturbing its eigenvalues [2, 155, 156], tensor invariants [3, 157] or the entire RST field [4]. One can also classify these stochastic models in terms of global and local perturbation (in space). For global approaches, such as in [22, 152, 153, 155], the magnitude of the perturbations in the closure coefficients, eigenvalues of RST, etc. is the

---

[1]Note that the notations in this chapter are slightly different from the previous chapters, as we follow the notations commonly used in the turbulence modeling literature.

This chapter is based on the article "Stochastic turbulence modeling in RANS simulations via Multilevel Monte Carlo method", *Preprint arXiv:1803.08864*, (2018) [27]. (Submitted for publication)

same throughout the flow domain. This translates to a low-dimensional UQ problem which can be efficiently handled by deterministic sampling techniques like stochastic collocation or just by simulating flows for limiting states to obtain a prediction interval. Since the error in closure models is not the same everywhere, global methods fail to capture the truth in general. On the other hand, local perturbation approaches may be effective due to high-dimensional parameterizations of the uncertainties. Some local methods already exist, such as the spatially varying marker functions proposed in [2, 158] or Gaussian random fields [4, 157, 159] as a measure of the local variation of the uncertainty. The main bottleneck hampering the development of these local models is the large cost of a forward uncertainty propagation stage.

The prime objective of this work is to provide a framework for developing a new class of high-dimensional stochastic RANS closures, that was until recently not viable (due to the cost of the propagation), but will be if the work required is within a constant, small factor of the cost of the fine-grid solution procedure. We achieve this using the multilevel Monte Carlo (MLMC) method [10, 29]. In previous works, the potential of the MLMC method has already been demonstrated in the context of the inviscid compressible flow in [160] for propagating lower-dimensional geometric and operational uncertainties. In the current work, we use two local stochastic models based on a random eddy viscosity and a random Reynolds stress tensor. The random eddy viscosity is obtained by perturbing the baseline eddy viscosity using Gaussian random fields with some prescribed spatial covariance. This stochastic model is applicable for the quantification of uncertainties arising due to imperfect closure constants. Similarly, the random RST is derived by perturbing the baseline RST. We utilize the algorithm proposed in [4] where the random RST is modeled by means of spatially correlated positive-definite random matrices. This approach is attractive as the random matrix is drawn from a set of positive-definite matrices which automatically guarantees realizable Reynolds stresses. Since the two stochastic extensions considered are high-dimensional in their random inputs, MC type methods are favorable owing to their dimension-independent convergence. For many UQ problems in fluid dynamics, the computational time and resources required to perform plain MC simulation are prohibitive. Standard MC methods may require thousands of CFD simulations on a fine computational mesh, before the statistical moments of the QoIs converge to some prescribed tolerance. The cost of the forward propagation can be drastically reduced by using the multilevel Monte Carlo method. In this work, we describe a standard MLMC method for efficient forward propagation of the uncertainty which is based on a hierarchy of pre-defined grids. For the proposed MLMC estimator, we show that the asymptotic cost does not deteriorate with an increase in the uncertain dimension and is controlled by the mesh convergence properties that further depend on the quality of the mesh and the discretization scheme used. For problems with a sufficiently fast decay of the numerical error, we demonstrate a cost scaling of $\mathcal{O}(\varepsilon^{-2})$ to achieve an error tolerance of ε. On the other hand, for problems with a slower error decay rate, we can attain an optimal MLMC estimator, in the sense that the cost grows at the same rate as the deterministic counterpart of the problem. The other motivation of this work is to show that the considered stochastic models can serve as an accurate Bayesian prior for calibration and data-assimilation involving turbulence models. Using numerical experiments, we show that the two models are sufficiently general and can reliably bound the possible flow behavior. Furthermore, the probability distribution of the random Reynolds stresses also satisfies the maximum entropy principle, a desirable property for a good prior.

The chapter is organized as follows. In Section 7.2 we briefly introduce the deterministic RANS equations and two standard deterministic turbulence models. Stochastic RANS models based on the random eddy viscosity and the random Reynolds stress are discussed in Section 7.3. In Section 7.4, numerical experiments are reported based on the flow over a periodic hill and fully developed turbulent flow in a square duct.

7.2. Deterministic RANS models

Direct numerical simulation of turbulent flow is highly expensive due to a large range of scales. Most engineering applications do not require details of these fine spatio-temporal features but only the effect of turbulence on the mean flow. A system of mean flow equations can be derived by Reynolds averaging, which consists of decomposing the flow into mean components, defined as an average over a large time period T, and fluctuations,

$$\overline{u_i} := \lim_{T \to \infty} \frac{1}{T} \int_0^T u_i \mathrm{d}t, \tag{7.1}$$

$$u_i' := u_i - \overline{u_i}, \tag{7.2}$$

respectively. The quantities $\overline{u_i}$ and u_i' are the mean and the fluctuating components of the instantaneous velocity u_i, respectively. Substituting (7.2) into the incompressible Navier-Stokes equation and averaging, results in the mean flow equation,

$$\rho(\overline{\mathbf{u}} \cdot \nabla)\overline{u}_i = -\frac{\partial \overline{p}}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\overline{R}_{ij} + R_{ij}\right), \quad i, j = 1, 2, 3. \tag{7.3}$$

The mean velocity vector is represented by $\overline{\mathbf{u}} = (\overline{u}_1, \overline{u}_2, \overline{u}_3)$, $\overline{p}$ is the time-averaged pressure field and ρ is the (constant) density. Here $\overline{R}_{ij} = \frac{1}{2}\rho\nu\left(\partial\overline{u}_i/\partial x_j + \partial\overline{u}_j/\partial x_i\right)$ denotes the mean stresses (tangential and normal) associated with the molecular viscosity ν. The mean flow is coupled to the turbulence by Reynolds stresses $R_{ij} = \rho\overline{u_i'u_j'}$. The Reynolds stress components R_{ij} appearing in (7.3) are unknown and are modeled using turbulence closure models that can be broadly categorized into Reynolds stress transport models and eddy viscosity models. The former models rely on an approximate set of stress transport equations to compute the Reynolds stress components. Although physically more stringent, stress transport models are not very popular in engineering practice as the discretizations of the coupled set of stress transport equations results in a numerically stiff system that is expensive to solve. On the other hand, linear eddy viscosity models are popular as they are significantly cheaper and perform reasonably well for a broad range of flows [148]. However, they are challenged by industrially relevant flows exhibiting separation, impinging, curvature, etc. These models are based on the *Boussinesq approximation*, which states that the Reynolds stresses are linearly related to the mean rate-of-strain as

$$-\overline{u_i'u_j'} \approx \nu_t\left(\frac{\partial\overline{u}_i}{\partial x_j} + \frac{\partial\overline{u}_j}{\partial x_i}\right) - \frac{2}{3}\delta_{ij}k, \tag{7.4}$$

where $k := \frac{1}{2}\overline{u_i'u_i'}$ is the turbulent kinetic energy, δ_{ij} is the Kronecker delta and ν_t is the eddy viscosity. On dimensional grounds the eddy-viscosity is a function of the turbulent velocity and length scale [161]. These quantities are commonly computed using two-equation turbulence models, such as $k - \epsilon$ or $k - \omega$, that are based on transport equations for k and for the turbulence-energy dissipation ϵ or the specific-dissipation ω. We use two popular turbulence models, the Launder-Sharma $k - \epsilon$ and a $k - \omega$ model. For both models, a generic k transport equation can be formulated with appropriate terms, listed in Table 7.1, as

$$\frac{\partial k}{\partial t} + \frac{\partial}{\partial x_j}\left[k\overline{u}_j - \left(\nu + \frac{\nu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right] = P - D, \quad \text{and} \quad P = R_{ij}\frac{\partial\overline{u}_i}{\partial x_j}. \tag{7.5}$$

The Launder-Sharma $k - \epsilon$ model is typically employed as a low-Reynolds number model. These kinds of models resolve the viscous part of the boundary layer with an appropriately refined

Term	Launder-Sharma $k - \epsilon$	$k - \omega$
D	$\epsilon + 2\nu \left(\dfrac{\partial \sqrt{k}}{\partial x_j} \right)^2$	$C_\mu \omega k$
ν_t	$C_\mu f_\mu \dfrac{k^2}{\epsilon}$	$\dfrac{k}{\omega}$
σ_k	1	2
C_μ	0.09	0.09

Table 7.1: Terms and coefficients in the k transport equation for two turbulence models.

mesh instead of utilizing wall functions [162]. Correct near wall behavior is achieved by damping functions for the eddy viscosity f_μ and the dissipation f_2 close to a wall. The equation for the dissipation ϵ reads

$$\frac{\partial \epsilon}{\partial t} + \frac{\partial}{\partial x_j} \left[\epsilon \overline{u}_j - \left(\nu + \frac{\nu_t}{\sigma_\epsilon} \right) \frac{\partial \epsilon}{\partial x_j} \right] = \left(C_{\epsilon_1} P - C_{\epsilon_2} f_2 \epsilon \right) \frac{\epsilon}{k} + 2\nu\nu_t \left(\frac{\partial^2 \overline{u}_i}{\partial x_j^2} \right)^2, \qquad \text{with} \qquad (7.6)$$

$$f_\mu = \exp \left[\frac{-3.4}{\left(1 + \frac{k^2}{50\nu\epsilon} \right)^2} \right], \; f_2 = 1 - 0.3 \exp \left[-\min \left(\left(\frac{k^2}{\nu\epsilon} \right)^2, 50 \right) \right], \qquad (7.7)$$

with $\sigma_\epsilon = 1.3$, $C_{\epsilon_1} = 1.44$, $C_{\epsilon_2} = 1.92$. The other model is the $k - \omega$ model [163], which uses a specific dissipation ω,

$$\frac{\partial \omega}{\partial t} + \frac{\partial}{\partial x_j} \left[\omega \overline{u}_j - \left(\nu + \frac{\nu_t}{\sigma_\omega} \right) \frac{\partial \epsilon}{\partial x_j} \right] = \gamma \frac{\omega}{k} P - \beta \omega^2, \qquad (7.8)$$

with $\sigma_\omega = 2$, $\gamma = 0.52$ and $\beta = 0.072$.

These two models form our baseline, to be perturbed in order to obtain stochastic RANS equations. However, the method proposed in this chapter is also applicable to other eddy viscosity models.

7.3. Stochastic RANS models

We now describe in detail the two stochastic models based on a perturbation of the baseline eddy viscosity field and the baseline Reynolds stress tensor field [4] originating from a deterministic eddy viscosity model. The former model is mathematically simple and is suitable for quantifying uncertainties that are introduced from a poor choice of RANS closure parameters to compute the eddy viscosity. The latter model is more advanced and is applicable to flows where the assumption of linear stress-strain relation is insufficient. When these models are combined with the RANS equations (7.3), we obtain so-called stochastic partial differential equations (SPDEs) that are solved using the MLMC method.

Before we describe the stochastic models, we clarify our setting. The RANS equations are defined in a bounded domain $\mathcal{D} \subset \mathbb{R}^d$ ($d = 1, 2, 3$). The complete probability space is denoted by $(\Omega, \mathbb{F}, \mathbb{P})$, where Ω is the sample space with σ-field $\mathbb{F}$ and probability measure $\mathbb{P}$. Furthermore, the zero-mean Gaussian random field will be denoted by $Z(\mathbf{x}, \omega)$, $\mathbf{x} \in \mathcal{D}$, $\omega \in \Omega$ with a specified positive-definite covariance kernel. We will work with a stationary anisotropic squared exponential

covariance model, given by

$$\mathbf{cov}(Z(\mathbf{x_1}, \cdot), Z(\mathbf{x_2}, \cdot)) = C(\mathbf{x_1}, \mathbf{x_2}) := \sigma_c^2 \exp\left(-\frac{(x_1 - x_2)^2}{l_x^2} - \frac{(y_1 - y_2)^2}{l_y^2} - \frac{(z_1 - z_2)^2}{l_z^2}\right), \quad (7.9)$$

where $C : \mathbb{R}^d \to \mathbb{R}_+$ with parameters σ_c^2 the marginal variance; l_x, l_y and l_z correlation lengths along the x, y and z directions, respectively. The algorithm for sampling $Z(\mathbf{x}, \omega)$ was described earlier in Section 3.6.

7.3.1   Random Eddy Viscosity (REV) model

RANS turbulence models rely on transport equations and a set of closure coefficients that are obtained from a calibration against DNS or experimental data. For a given turbulence model, a closure coefficient take different values when calibrated against different types of flow [148]. Since the model prediction is strongly influenced by the value of the closure coefficients, a common practice is to propagate a joint probability distribution of these closure parameters to obtain uncertainty bounds of the QoIs, see for e.g., [22, 152, 154]. These approaches indirectly lead to a globally perturbed eddy viscosity field. Here, one must take into account the fact that the Boussinesq assumption (7.4) is in the general case only locally imperfect. Therefore, methods that allow direct local perturbations of the baseline eddy viscosity fields can be effective. A convenient way to achieve this local perturbation is by the means of Gaussian random fields with some prescribed covariance model. Depending on the problem, a covariance model can be designed which induces a high-variability locally in ν_t; around regions where eddy viscosity models are expected to perform poorly. The samples of the random eddy viscosity field $\nu_t(\mathbf{x}, \omega)$ can be obtained by perturbing the baseline eddy viscosity field which we now denote by $\nu_t^{(bl)}(\mathbf{x})$ with the Gaussian random field,

$$\log \nu_t(\mathbf{x}, \omega) = \log \nu_t^{(bl)}(\mathbf{x}) + Z(\mathbf{x}, \omega), \quad (7.10)$$

where ω denotes the random event in the stochastic domain Ω. The mean field $\nu_t^{(bl)}$ is obtained from a converged deterministic RANS simulation and is based on a baseline turbulence model, or from an average of eddy viscosities obtained from different turbulent models. The above relation is the simplest multiplicative model, $\nu_t(\mathbf{x}, \omega) = \nu_t^{(bl)}(\mathbf{x})e^{Z(\mathbf{x},\omega)}$, that is able to impose positivity of random eddy viscosity samples and values close to zero near the wall region. We point out that Dow and Wang [159, 164] also explored Gaussian random fields to obtain uncertainty bounds in the mean flow. In their approach the variability of the Gaussian process was based on the discrepancy between eddy viscosities obtained from the DNS data (known as the "true" eddy viscosity) and those predicted by any turbulence model.

With the random eddy viscosity, we obtain the following SPDE:

$$\rho(\overline{\mathbf{u}} \cdot \nabla)\overline{u}_i = -\frac{\partial \overline{p}^*}{\partial x_i} + \frac{\partial}{\partial x_j}\left[(\nu + \nu_t(\omega))\left(\frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i}\right)\right], \quad (7.11)$$

where $\overline{p}^* := \overline{p} - \frac{2}{3}k$. Recall that the above SPDE can be used for quantifying uncertainties due to inconsistencies in the closure parameters of the baseline model and also provide a way to account for the effect of local variations of these parameters in the flow, unlike [22, 152, 154]. However, this stochastic model still inherits drawbacks from the Boussinesq hypothesis and is inadequate for quantifying uncertainties associated with turbulence anisotropy. For instance, the occurrence of secondary flows as a result of normal stress imbalance (e.g. in a square duct) will remain undetected. Therefore, a more generic stochastic model is also discussed, that involves the injection of uncertainties directly into the baseline Reynolds stress tensor.

7.3.2   Random Reynolds Stress Tensor (RRST) model

The RRST model stems from the work by Soize in [165–168] who developed non-parametric prob-
abilistic approaches based on random matrix theory to quantify modeling uncertainties in com-
putational mechanics problems. Soize derived the maximum entropy probability distribution for
symmetric positive-definite (SPD) real matrices with a given mean and variance (also known as
the *dispersion* parameter, δ) along with a Monte Carlo sampling method. These results with slight
modifications can be utilized for the sampling of random Reynolds stress tensors (as physically
realizable RSTs are symmetric positive semi-definite matrices). Xiao and coworkers in [4] further
extended this approach to incorporate spatial correlation in the Reynolds stress components by the
means of Gaussian random fields with a prescribed covariance function. We now briefly outline
sampling algorithms for a random SPD matrix that will be utilized later to describe the sampling of
the random Reynolds stress tensor fields. We closely follow the description from the original papers
[4, 168, 169].

Sampling random SPD matrices

We denote by $\mathbb{M}_d^{+0}(\mathbb{R})$ and $\mathbb{M}_d^{+}(\mathbb{R})$ the set of all $d \times d$ symmetric positive semi-definite and symmetric
positive-definite matrices with real entries, respectively. Given a baseline matrix $\mathbf{R}_{(bl)} \in \mathbb{M}_d^{+}(\mathbb{R})$,
we wish to sample random matrices $\mathbf{R} \in \mathbb{M}_d^{+}(\mathbb{R})$, such that $\mathbb{E}[\mathbf{R}] = \mathbf{R}_{(bl)}$. The sampling of $\mathbf{R}$ can be
achieved using a *normalized* random SPD matrix $\mathbf{G} \in \mathbb{M}_d^{+}(\mathbb{R})$ with mean $\mathbf{I}_d$ (identity), i.e. $\mathbb{E}[\mathbf{G}] = \mathbf{I}_d$
and the variance parameterized with a dispersion parameter $\delta > 0$ defined as

$$\delta = \sqrt{\frac{1}{d}\mathbb{E}\big[||\mathbf{G} - \mathbf{I}_d||_F^2\big]}, \tag{7.12}$$

where $|| \cdot ||_F$ is the Frobenius norm. A first step is to utilize the Cholesky decomposition $\mathbf{G} = \mathbf{U}^T\mathbf{U}$,
where $\mathbf{U}$ is an upper-triangular matrix with positive diagonal entries. Now, the assembly of the
random matrix $\mathbf{G}$ boils down to sampling the six entries of $\mathbf{U}$. The non-diagonal entries of $\mathbf{U}$ are
sampled by means of

$$U_{ij} = \frac{\delta}{\sqrt{d+1}}\xi_{ij}, \quad \text{for} \quad i > j, \quad \xi_{ij} \sim \mathcal{N}(0,1). \tag{7.13}$$

The diagonal entries are sampled as

$$U_{ii} = \frac{\delta}{\sqrt{d+1}}\sqrt{2y_i}, \quad \text{for} \quad i = 1,2,3, \tag{7.14}$$

where $y_i > 0$ is a sample from the gamma distribution $\mathbf{\Gamma}(\kappa_i, 1)$ with shape parameter κ_i and scaling
parameter 1, i.e.

$$y_i \sim \mathbf{\Gamma}(\kappa_i, 1) \quad \text{with} \quad \kappa_i = \frac{d+1}{2\delta^2} + \frac{1-i}{2}. \tag{7.15}$$

The gamma probability density function f_Y is given by:

$$f_Y(y_i) = \frac{y_i^{\kappa_i - 1}\exp(-y_i)}{\Gamma(\kappa_i)} \quad \text{with} \quad \kappa_i = \frac{d+1}{2\delta^2} + \frac{1-i}{2}, \tag{7.16}$$

where $\Gamma(\cdot)$ is the standard gamma function. For different diagonal terms, $y_i(\mathbf{x}, \cdot)$ will have different
marginal PDFs depending on the shape parameter κ_i. Using $\mathbf{G}$, one can obtain the random matrix
$\mathbf{R}$ with mean $\mathbf{R}_{(bl)}$ as:

$$\mathbf{R} = \mathbf{U}_{(bl)}^T\mathbf{G}\mathbf{U}_{(bl)}, \tag{7.17}$$

where $\mathbf{U}_{(bl)}$ is an upper-triangular matrix with positive diagonal entries obtained via the Cholesky factorization of the baseline RST $\mathbf{R}_{(bl)} = \mathbf{U}_{(bl)}^T \mathbf{U}_{(bl)}$. Assuming $\mathbf{R}_{(bl)}$ to be positive-definite, the factorization yields a unique matrix $\mathbf{U}_{(bl)}$. Note that in practice $\mathbf{R}_{(bl)}$ is symmetric positive semi-definite, belonging to $\mathbb{M}_d^{+0}(\mathbb{R})$. The RSTs with zero eigenvalues i.e. $\mathbf{R}_{(bl)} \in \mathbb{M}_d^{+0}(\mathbb{R}) \backslash \mathbb{M}_d^+(\mathbb{R})$ are only encountered when $\det(\mathbf{R}_{(bl)}) = 0$, corresponding to the 2-component turbulence limit [148]. However, adding an arbitrarily small number to the diagonal will make this tensor symmetric positive-definite. We also point out that, to maintain positive-definiteness of $\mathbf{G}$, the dispersion parameter δ should be chosen such that $0 < \delta < \sqrt{(d+1)(d+5)^{-1}}$, see [167] for details. Thus, for $d = 3$, we find the constraint $0 < \delta < 1/\sqrt{2}$.

Sampling the random tensor field

The sampling algorithm for SPD matrices can be extended to sample spatially correlated tensor fields. We follow a similar procedure as described in the preceding section but now the entries of the upper-triangular matrix $\mathbf{U}$ are correlated in space. We describe the necessary algorithmic modifications needed to sample these random RST fields.

Let the random RST at any point be denoted by $\mathbf{R}(\mathbf{x}, \omega) = \mathbf{R}$, the deterministic baseline Reynolds stress tensor field by $\mathbf{R}_{(bl)}(\mathbf{x}) = \mathbf{R}_{(bl)}$ and a spatially varying dispersion field by $\delta(\mathbf{x})$. Furthermore, the entries of the random upper-triangular matrix, $\mathbf{U}(\mathbf{x}, \omega) = \mathbf{U}$, are spatially correlated as:

$$\mathbf{cov}\{U_{ij}(\mathbf{x_1}, \cdot), U_{ij}(\mathbf{x_2}, \cdot)\} = C(\mathbf{x_1}, \mathbf{x_2}), \qquad i > j, \tag{7.18}$$

$$\mathbf{cov}\{U_{ii}^2(\mathbf{x_1}, \cdot), U_{ii}^2(\mathbf{x_2}, \cdot)\} = C(\mathbf{x_1}, \mathbf{x_2}), \qquad i = j. \tag{7.19}$$

As suggested in [4], we also consider a squared-exponential covariance function for both off-diagonal and for the square of the diagonal terms. Other covariance models, for instance, a periodic or an exponential covariance can also be utilized. For the sake of simplicity, we use $C(\mathbf{x_1}, \mathbf{x_2})$ defined in (7.9) but with $\sigma_c^2 = 1$. Now, the random tensor field $\mathbf{R}$ is assembled using six independent random fields: $U_{11}(\mathbf{x}, \omega)$, $U_{12}(\mathbf{x}, \omega)$, $U_{13}(\mathbf{x}, \omega)$, $U_{22}(\mathbf{x}, \omega)$, $U_{23}(\mathbf{x}, \omega)$ and $U_{33}(\mathbf{x}, \omega)$. The off-diagonal fields are computed as:

$$U_{ij}(\mathbf{x}, \omega) = \frac{\delta(\mathbf{x})}{\sqrt{d+1}} Z_{ij}(\mathbf{x}, \omega), \quad \text{for} \quad i > j, \quad Z_{ij} \sim \mathcal{N}(0, C). \tag{7.20}$$

The Gaussian random fields Z_{ij} can be generated using KL expansion, as described in (3.42). Similar to (7.14), the diagonal elements are obtained as:

$$U_{ii}(\mathbf{x}, \omega) = \frac{\delta(\mathbf{x})}{\sqrt{d+1}} \sqrt{2y_i(\mathbf{x}, \omega)}, \quad \text{for} \quad i = 1, 2, 3, \tag{7.21}$$

where $y_i(\mathbf{x}, \omega) > 0$ denotes a random field with gamma marginal distribution $\Gamma(\kappa_i(\mathbf{x}), 1)$ and covariance defined in (7.19). Now, the marginal gamma PDF in (7.16) is modified to incorporate spatial dependence by $\delta(\mathbf{x})$ as

$$f_Y(y_i(\mathbf{x}, \cdot)) = \frac{y_i(\mathbf{x}, \cdot)^{(\kappa_i(\mathbf{x})-1)} \exp(-y_i(\mathbf{x}, \cdot))}{\Gamma(\kappa_i(\mathbf{x}))}, \quad \text{with} \quad \kappa_i(\mathbf{x}) = \frac{(d+1)}{2\delta(\mathbf{x})^2} + \frac{(1-i)}{2}. \tag{7.22}$$

We use the same technique as in Section 6.3.1 to sample non-Gaussian random field. The method relies on a generalised Polynomial Chaos (gPC) expansion to approximate non-Gaussian field in terms of a weighted combination of Hermite orthogonal polynomials of a standard Gaussian field i.e.

$$Y \approx \sum_{n=1}^{N_{PC}} \mathrm{w}_n \mathcal{H}_n(Z), \tag{7.23}$$

where Y represents a spatially correlated gamma random field, N_{PC} is the order of the expansion and $\mathcal{H}_n(Z)$ is the Hermite polynomial in Z of order n with weight w_n. Given the orthogonality of Hermite polynomials with respect to the Gaussian measure, we can evaluate the weights as:

$$\mathrm{w}_n = \frac{\mathbb{E}[Y\mathcal{H}_n(Z)]}{\mathbb{E}[\mathcal{H}_n(Z)^2]}. \tag{7.24}$$

Here the expectation in the denominator has an analytic expression but the expectation in the numerator is not well-defined as the dependence between Y and Z is unknown. Since the distribution of Y is available, one can exploit the fact that $Y = F_Y^{-1}(F_Z(Z))$ and reformulate the numerator in (7.24) as

$$\mathbb{E}[Y\mathcal{H}_n(Z)] = \int_{-\infty}^{\infty} F_Y^{-1}[F_Z(z)]\mathcal{H}_n(z)\mathrm{d}F_Z(z), \tag{7.25}$$

where $F_Y(y) = \mathbb{P}\mathrm{rob}(Y \leq y)$ is the cumulative distribution for a gamma random variable Y and F_Y^{-1} represents its inverse. Similarly, $F_Z(z) = \mathbb{P}\mathrm{rob}(Z \leq z)$ is the cumulative distribution for a standard Gaussian random variable Z. Now, the integral (7.25) can be numerically computed using any conventional integration technique. With the above weights, the gPC expansion in (7.23) converges to Y in *weak sense* (convergence in probability distribution) [7, 170]. Note that F_Y should be appropriately modified according to (7.22) to incorporate the spatial dependence in the marginal gamma PDF. It is also pointed out that for a spatially varying dispersion $\delta(\mathbf{x})$ the weights will differ at different spatial locations.

A few remarks are in order. The mean RST field $\mathbf{R}_{(bl)}$ can be directly obtained from the baseline RANS simulation. Also, the value of the dispersion field can be based on expert knowledge and can be set to a large value at locations with high uncertainty. However, to obtain a positive-definite Reynolds stress tensor at each point the dispersion should again be chosen such that $0 < \delta(\mathbf{x}) < \sqrt{(d+1)(d+5)^{-1}}$.

Using the random Reynolds stress tensor, we can define the stochastic mean flow equation, as follows:

$$\rho(\overline{\mathbf{u}} \cdot \nabla)\overline{u}_i = -\frac{\partial \overline{p}}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\overline{R}_{ij} + R_{ij}(\omega)\right), \tag{7.26}$$

where $\overline{R}_{ij}$ represents mean stress, as defined for the PDE (7.3) and $R_{ij}(\omega)$ represents components of the random tensor field $\mathbf{R}$. In this stochastic model the isotropic eddy viscosity (Boussinesq) assumption is clearly avoided. Furthermore, this model allows us to accommodate different covariance structures for different Reynolds stress components, and thus can represent strongly anisotropic turbulence. We would like to emphasize that the above SPDE is more general than in (7.11) as the above formulation allows us to incorporate at most six random fields for each Reynolds stress component and may result in an extremely high-dimensional UQ problem.

7.3.3   MLMC-RANS implementation

The MLMC-RANS framework is developed in MATLAB and interacts with the OpenFOAM (Open source Field Operation And Manipulation) CFD package [171]. MATLAB based programs are responsible for the generation of random inputs (eddy viscosity fields and Reynolds stress tensors), invoking OpenFOAM with random inputs, the collection of samples of the QoI and post-processing. Within OpenFOAM, schemes for the computation of the gradients and divergence are based on second-order finite volume (FV) approximations. The baseline solution of the turbulence models is obtained using the simpleFoam solver [172] available in OpenFOAM, and to propagate the random eddy viscosity and random Reynolds stresses different solvers were implemented for the stochastic momentum equations (7.11) and (7.26), respectively.

While the propagation of random eddy viscosity is straightforward and doesn't require modification of the solver in general, the propagation of random Reynolds stresses is numerically more challenging. To achieve a numerically stable performance of the solver, we adopt a blending of the random Reynolds stress, which we wish to propagate, and a contribution based on the Boussinesq assumption [173]. While the latter alters the propagated effective Reynolds stress, it promotes numerical convergence of the solver. The momentum equation (7.26) is modified accordingly,

$$\rho(\overline{\mathbf{u}} \cdot \nabla)\overline{u}_i = -\frac{\partial \overline{p}}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\overline{R}_{ij} + (1 - \xi)R_{ij}^{(bl)} + \xi R_{ij}(\omega)\right), \tag{7.27}$$

in which the linear eddy viscosity contribution $R_{ij}^{(bl)}$ is given in (7.4). The production of turbulent kinetic energy is modified accordingly. The blending parameter $\xi \in [0, 1]$ quantifies the amount of $R_{ij}^{(bl)}$ to increase numerical stability. For $\xi = 1$, we achieve the full propagation of the random tensor field. This is possible in case of simpler flows, for e.g., flow in a square duct. Also, the value of ξ is linearly increased with the number of iterations (ramping) to a constant value. Note that a value of $\xi < 1$ indirectly corresponds to lower variance, than specified for a given dispersion δ.

To facilitate the analysis, our implementation of the MLMC method is based on a pre-defined geometric hierarchy of meshes such that the largest cell width follows $h_{\ell-1} \approx 2h_\ell$. In general, an MLMC estimator can be constructed with any hierarchy for which the accuracy and cost increase with the levels. The quality of the mesh at any given MLMC level ℓ is assessed using the dimensionless wall distance, defined as $y_\ell^{+1} = h_\ell^{cc} u_\ell^* / \nu$ where h_ℓ^{cc} denotes the distance of the cell-centers adjacent to the wall, u_ℓ^* is the friction velocity defined as $u_\ell^* = \sqrt{\tau_\ell^w / \rho}$ and the wall shear stress $\tau_\ell^w = \mu(\partial u/\partial y)_{y=0}$. The standard notation ν and μ is used for kinematic and dynamic viscosities, respectively. For resolving the viscous sublayer, the y_ℓ^{+1} value should be less than one, however, this criterion can be relaxed for coarser levels in the MLMC hierarchy provided that the RANS solution results in a meaningful flow field. Furthermore, we check that the level-dependent variance should be strictly less than the pure sample variance of the quantity of interest, i.e., $\|\mathcal{V}[u_\ell - u_{\ell-1}]\|_{L^2(\mathcal{D})} < \|\mathcal{V}[u_\ell]\|_{L^2(\mathcal{D})}$. Violation of this condition may result in an MLMC estimator which is more expensive than a standard MC estimator.

As this work involves stationary covariance models, we use a spectral generator for the fast sampling of the Gaussian random fields, see Section 3.6 for details. It is pointed out that with this algorithm the computational cost of sampling a random field is of the order $\mathcal{O}(M_\ell \log M_\ell)$, where M_ℓ is the number of mesh points on any level ℓ and is negligible compared to the cost of one CFD solve at that level. Additionally, the random fields generated using spectral methods are exact on the sampling mesh. In the case of the KL expansion based sampling, one needs to quantify the error incurred due to the truncation of the eigenmodes.

7.4. Numerical experiments

We use two test problems, a fully developed turbulent flow in a square duct and a flow over a periodic hill, to study the performance of the MLMC method. A bulk Reynolds number $Re = 1100$ is considered for the square duct flow with benchmark data available from Huser et al. (1993) [174]. This problem has become a standard test case to demonstrate the inability of linear eddy viscosity models to predict the secondary flows that arise from the normal stress imbalance. Linear eddy viscosity models assume equal normal stresses and completely fail to predict secondary flow features, resulting in parallel flow. We only employ the random Reynolds stress model for this test case as the random eddy viscosity model suffers from the same drawback as the deterministic linear eddy viscosity model and fails to produce any secondary flows. For the periodic hill problem, we use $Re = 2800$ with the DNS data from Breuer et al. (2009) [175]. This is a complex benchmarking
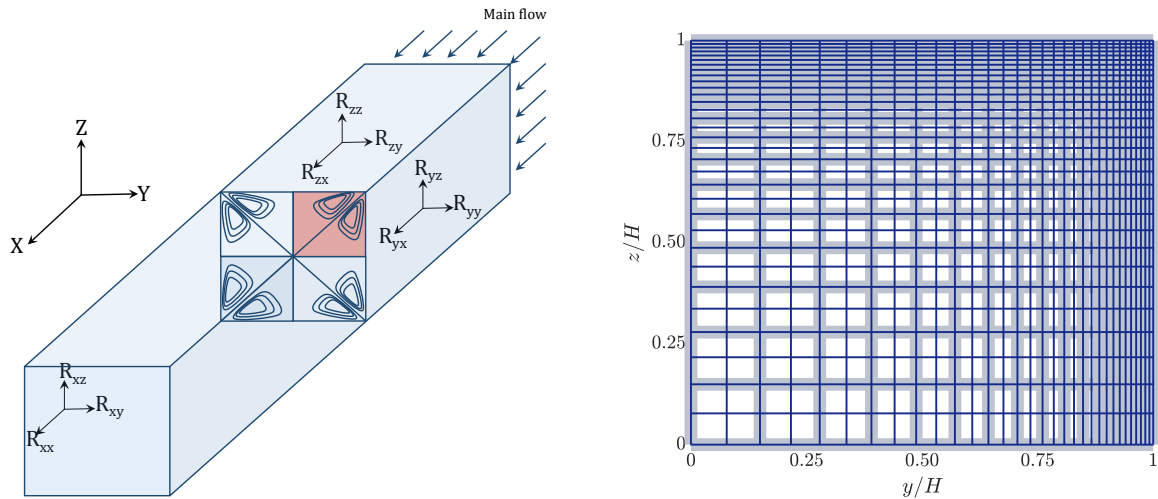
Figure 7.1: Schematic representation of time-averaged flow in a square duct (left) showing the 8-vortex pattern with each quadrant exhibiting vortices of alternating sign. (Right) Nested meshes $\ell = 0$ (light blue) and $\ell = 1$ (dark blue) used to simulate the flow in the top-right quadrant, grading near the walls.

test problem, offering a number of flow features such as anisotropy, strong streamline curvature, a recirculating zone and free shear layer, that are challenging for RANS turbulence models. Both stochastic models are analyzed for the periodic hill flow.

7.4.1 Flow in a square duct

A schematic representation of the square duct flow is presented in Figure 7.1 (left) showing the eight-vortex pattern with counter-rotating vortices in each quadrant. Due to symmetry, we choose to simulate the flow only for the top-right quadrant on a domain of size $[0, H] \times [0, H]$, where $H = 1$ is the half-height of the square duct. We use a separate grid hierarchy for the OpenFOAM simulations and for sampling the random fields, denoted by OF and RF meshes, respectively, with specifications listed in Table 7.2. For the OF meshes, each grid level is graded with finer cells along the top and right walls to resolve boundary layers, see Figure 7.1 (right). In the case of RF meshes, the random fields are first sampled on a uniform Cartesian mesh in the domain $[0, 1]^2$ and are then interpolated to the cell-centers of the RANS simulation mesh. The cell-width of the RF mesh is finer than the largest cell-width of the OF mesh, h_ℓ. Further, a second-order accurate interpolation scheme is used to ensure that the interpolation error is sufficiently smaller than the discretization bias. The CPU times on a serial machine required to obtain one sample on each level is also provided in Table 7.2. For the considered combination of numerical schemes, the CPU times scale roughly as $\mathcal{O}(h_\ell^{-3})$ (in other words, $\gamma \approx 3$). This is due to the fact that the convergence rate of the solver deteriorates with grid refinement, therefore, the number of iterations required to reach a fixed residual tolerance also grows with levels. Additionally, the residual tolerance also needs to be reduced with grid refinement in order to obtain a converged solution up to the discretization accuracy, and on the finest levels, one sample takes about eight CPU hours to obtain a residual reduction of $\mathcal{O}(10^{-8})$.

MLMC with RRST model

We begin by analyzing the statistics of the random Reynolds stress tensors for two sets of parameters (Case 1 and Case 2) as specified in Table 7.3. Here, we can regard Case 1 as an "easy" parameter set, with a low dispersion and large correlation lengths and Case 2 as "more complex" with a large dispersion and small correlation lengths. For both cases a 5th order gPC expansion is used such that errors in approximating the random field are negligible compared to the discretization and

Level (ℓ)	OF mesh	h_ℓ	RF mesh	CPU time (sec)
0	16×16	0.16	8×8	0.24×10^2
1	32×32	0.08	16×16	0.68×10^2
2	64×64	0.04	32×32	4.20×10^2
3	128×128	0.02	64×64	2.86×10^3
4	256×256	0.01	128×128	2.93×10^4

Table 7.2: Specifications of the MLMC grid hierarchy for the square duct test case. "OF mesh" denotes the simulation mesh in OpenFOAM and "RF mesh" the grid used for the generation of the random Reynolds stress tensor. CPU time is the total time for one sample.

sampling errors. In this work, we will only consider cases with a constant dispersion, but a more general approach can be based on a spatially varying dispersion based on available data and expert knowledge as in [4, 157]. For both cases, a full propagation of the random Reynolds stress (i.e. $\xi = 1$) is considered.

In Figure 7.2, we present examples of the first three Reynolds stress components, R_{11}, R_{12}, R_{13}, generated using the two parameter sets along with the baseline Reynolds stress tensors $\mathbf{R}_{(bl)}$ (derived from the $k - \omega$ model). Firstly, we verify the constraint $\mathbb{E}[\mathbf{R}] = \mathbf{R}_{(bl)}$ by computing the empirical probability distribution using around 1.6×10^4 samples on the coarsest 16×16 grid level. The empirical PDFs for the first three components of the Reynolds stress at a location inside one of the vortices $(y/H, z/H) = (0.52, 0.21)$ are presented in Figure 7.3 for the two cases. The PDFs of other components of the Reynolds stress tensor exhibit similar behaviour, and are omitted. We observe that the sample mean is very close to the baseline value and for Case 2, due to a larger δ, a slight deviation ($\sim 5 \times 10^{-4}$) from the baseline is observed, consistent with the sampling error. The state of the anisotropy resulting from the samples of the random Reynolds stresses is visualized using the barycentric triangle [176] in Figure 7.4. Again the probability density contours are based on 1.6×10^4 samples at location $(y/H, z/H) = (0.52, 0.21)$ for each case. The procedure to construct these contours is explained in Appendix 7.A. We observe that the distance between the state of anisotropy obtained from the baseline simulation and the sample mean is sensitive to the dispersion parameter. For a larger dispersion, many samples fall away from the baseline state but due to the positive-definite constraint they are restricted until the edges of the barycentric triangle. Thus, the sample mean is located far from the baseline anisotropy state, see [4] for details. The effect of this constraint is mild for a smaller dispersion and the mean anisotropy state is very close to the baseline.

Parameter	Description	Case 1	Case 2
$l_y/H, l_z/H$	Correlation length along y/z-direction	2	1
σ_c^2	Variance of log-normal random field	1	1
$\delta(\mathbf{x})$	Dispersion parameter	0.1	0.4
N_{PC}	Order of polynomial chaos expansion	5	5
ξ	Blending factor	1	1

Table 7.3: Parameter sets to generate the random Reynolds stress tensor for the square duct flow.

We begin by studying the FV error convergence for Case 1 and Case 2. We will only consider the u and v components of the velocity as w has similar characteristics as v. In Figure 7.5, we show the relative error $||u_\ell - u_{\ell-1}||_{L^2(\Omega, \mathcal{D})}$ along with the FV errors from the deterministic RANS simulations
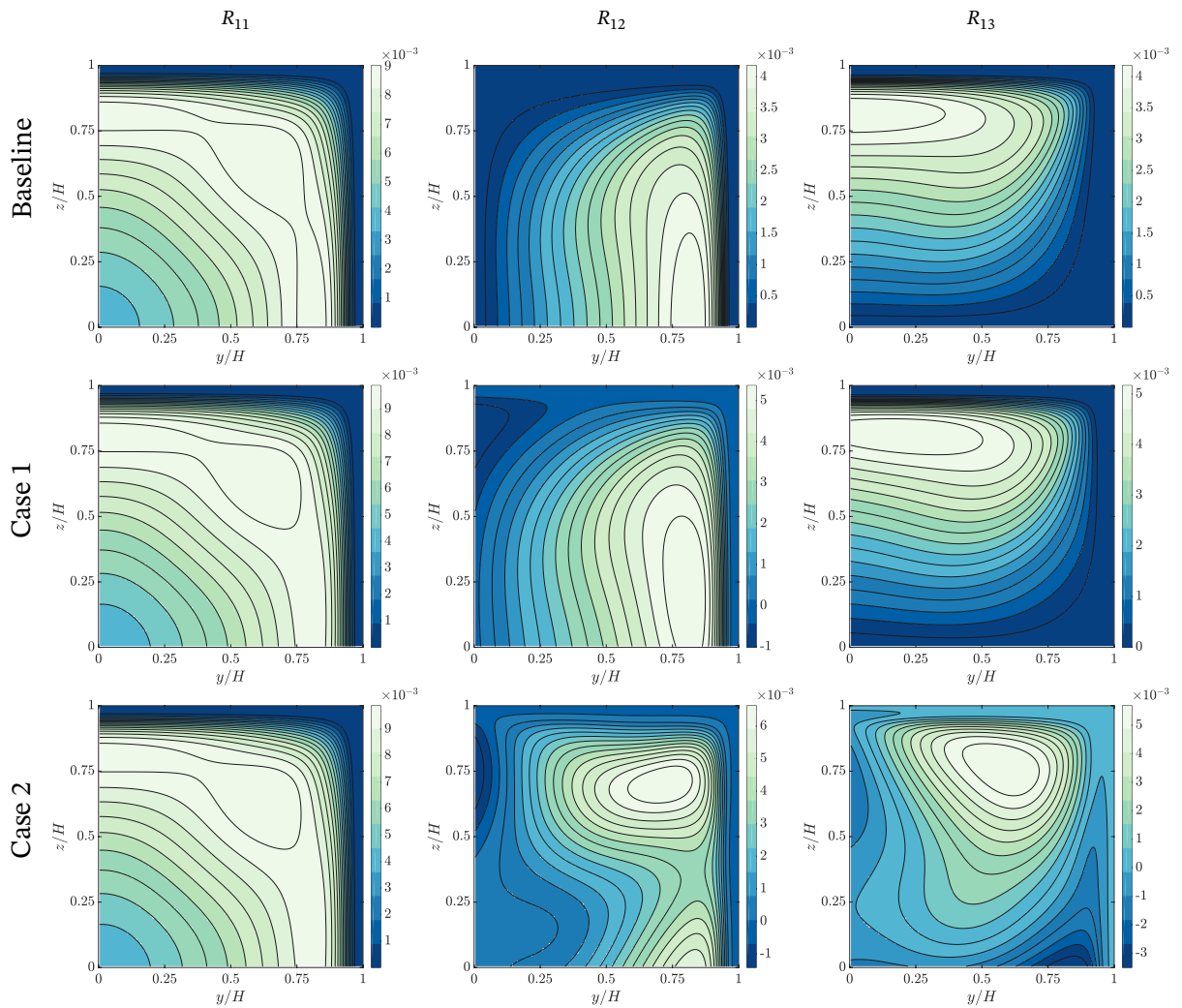
Figure 7.2: Reynolds stress components, R_{11}, R_{12}, R_{13}, obtained from the baseline $k - \omega$ model (top row) and an example of perturbed random Reynolds stresses generated from Case 1 (middle row) and Case 2 (bottom row).

(based on the $k - \omega$ model) plotted against the maximum cell width h_ℓ. The relative error for v is also presented in Figure 7.5 (right). As the deterministic RANS simulation predicts $v = 0$, we again use the deterministic error in u for comparison of the FV convergence rates. These relative errors are computed with a sufficient number of samples such that sampling errors on each level are less than the FV bias. We observe a convergence of $\mathcal{O}(h_\ell^{1.5})$ (rounded to one decimal place) for the deterministic simulations and further note that the stochastic version of the FV error also decays at a similar rate. Here, we remark that although we use second-order accurate schemes, a slightly slower error convergence is obtained, most likely due to the non-uniformity of the meshes used. Also, for the deterministic simulations on the finest 256×256 grid, OpenFOAM has convergence issues. Interestingly, this is not observed for the stochastic simulations. Further, due to a higher value of the dispersion parameter δ for Case 2, compared to Case 1, we see a larger absolute numerical error, but it decays at a similar rate. These plots are important in order to determine the number of levels that should be included in the MLMC hierarchy to reduce the RMSE to a given tolerance of ε. For the standard Monte Carlo simulation, the error associated with a particular mesh is utilized to determine the number of samples needed on that mesh, to equilibrate the sampling error with the discretization error.

The convergence of the level-dependent variance $\|\mathcal{V}[\cdot]\|_{L^2(\mathcal{D})}$ is shown in Figure 7.6. For refer-
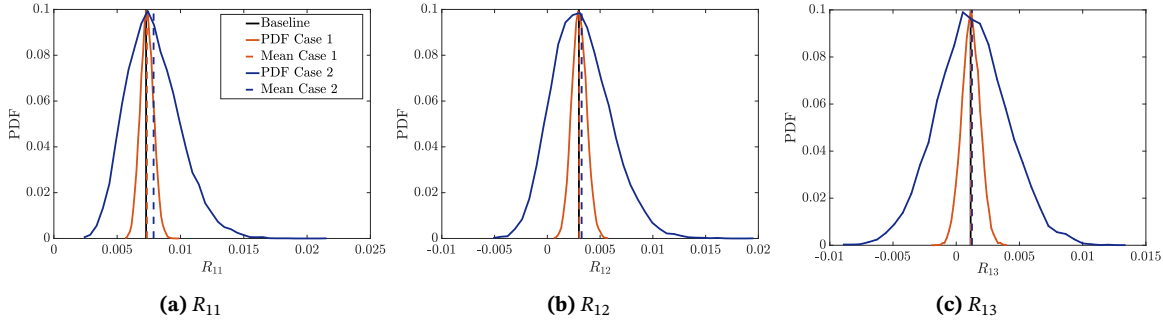
Figure 7.3: Empirical PDF of the Reynolds stress components at location $(y/H, z.H) = (0.52, 0.21)$ for Case 1 ($\delta = 0.1$) and Case 2 ($\delta = 0.4$). For the diagonal component R_{11}, a gamma marginal distribution is obtained and for the off-diagonal components R_{12}, R_{13}, Gaussian distributions are observed.
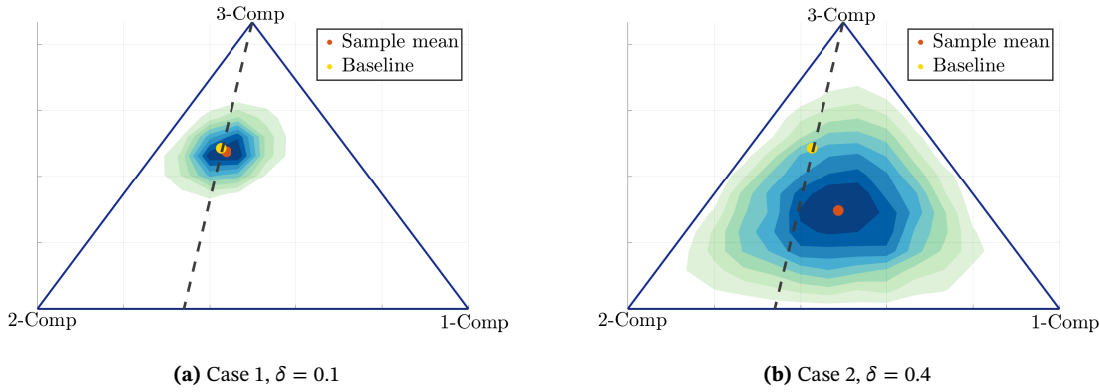
Figure 7.4: Probability density contours for random Reynolds stresses at location $(y/H, z/H) = (0.52, 0.21)$ projected to the barycentric triangle based on 1.6×10^4 samples on $\ell = 0$.

ence, an $\mathcal{O}(h_\ell^3)$ convergence line is plotted to emphasize $\beta \approx 2\alpha$. The significance of these plots is that they can be used to assess the sampling variance at different levels and extract the rate β used to determine the MLMC sample sequence in the formula (2.29). We observe a higher variance for larger dispersion from Case 2 compared to Case 1, as expected. We point out that the above convergence study can be quite expensive, as many samples over all the levels are needed to obtain accurate estimates of the MLMC rates α, β. The purpose of the above analysis is to (i) demonstrate that the FV error decay rate extracted from the deterministic solves can be an accurate estimate of the rate α and (ii) verify that the assumption $\beta = 2\alpha$ holds. With a fixed β, we can obtain the number of samples on all MLMC levels in advance and can efficiently distribute the work on a computing cluster. Alternatively, one can also implement the standard MLMC algorithm [10, 28] which adaptively computes and refines these parameters along with the number of samples on each level until a prescribed tolerance is achieved. Note that for such algorithms, optimal load balancing is non-trivial as the number of samples predicted on the different levels after every refinement stage varies.

Next, we compare the accuracy and computational cost of the MLMC and MC estimators to compute the mean and variance. For this analysis the in-plane velocity v is chosen as the quantity of interest. To measure the accuracy, we rely on the following relative error measure [30, 32]:

$$\varepsilon_{rel} := \frac{||\mathcal{E}_{ref}[v] - \mathcal{E}_L^{ML}[v_L]||_{L^2(\mathcal{D}_L)}}{||\mathcal{E}_{ref}[v]||_{L^2(\mathcal{D}_L)}}. \tag{7.28}$$

Here, $\mathcal{E}_L^{ML}[v_L]$ can be replaced by the standard MC estimator $\mathcal{E}_N^{MC}[v_h]$. Analogously, the relative
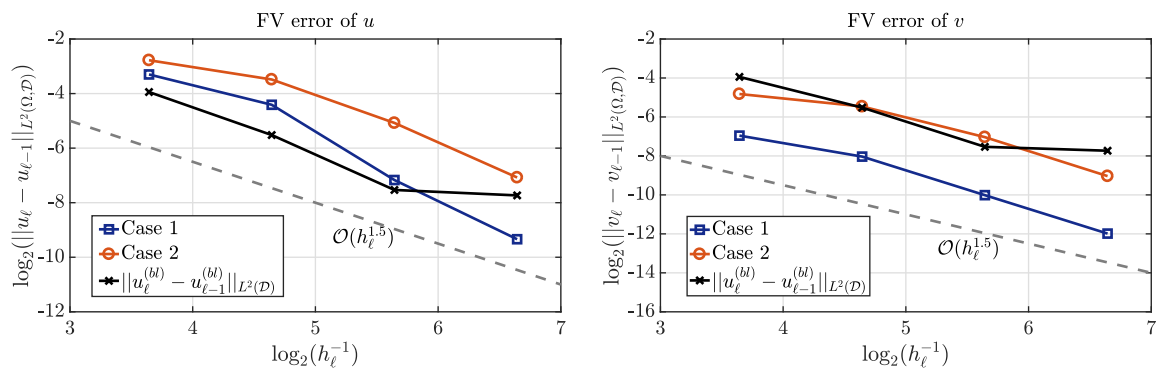
Figure 7.5: Convergence of the FV error with levels along with the error in baseline solution of u. The dotted line denotes the empirical convergence rate of baseline RANS simulations.
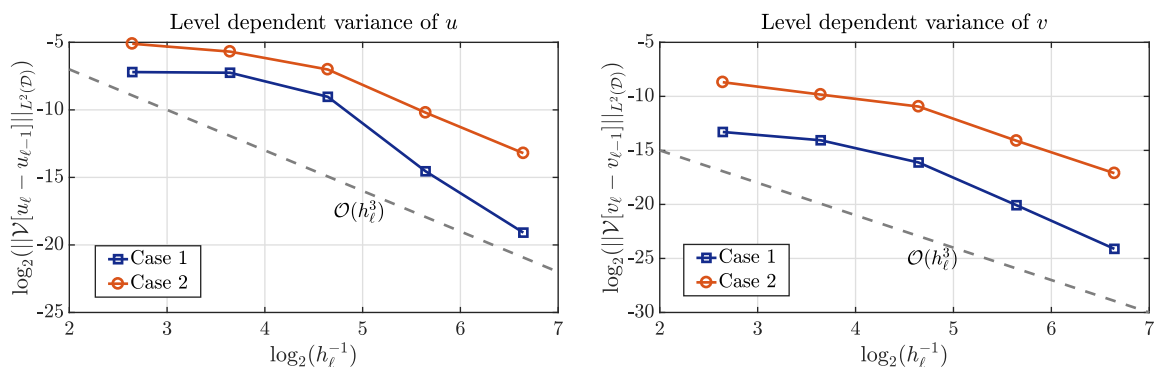

Figure 7.6: Convergence of the level-dependent variance with grid refinement. The dotted line depicts an $\mathcal{O}(h_\ell^3)$ convergence.

errors in the variance estimators $\mathcal{V}_N^{MC}$ and $\mathcal{V}_L^{ML}$ are also computed. For the MLMC estimator, we compute the mean and variance for different h_L (or h for the standard MC). These experiments are conducted 16 times to eliminate statistical fluctuations and the mean relative error $\bar{\varepsilon}_{rel}$ is reported. The reference solutions for the expected value $\mathcal{E}_{ref}[v]$ and the variance $\mathcal{V}_{ref}[v]$ are computed using the 5-level MLMC estimator. Reference solutions will be discussed in detail later on.

Based on the deterministic FV error convergence study, we fix $\alpha = 1.5$ and $\beta = 2\alpha = 3$ and $\gamma = 3$ (see Table 7.2). Thus, for the MLMC estimator, we get a sample sequence $N_\ell = N_L 2^{3(L-\ell)}$ based on the formula (2.29). Note that we have $\beta \approx \gamma$ and therefore we can obtain an MLMC estimator for which all levels contribute equally in terms of the cost, see (2.27). As mentioned earlier, the number of samples on the finest level N_L is a free parameter and should be set to a small value. For all experiments, we use $N_L = 8$. In Table 7.4, we list the number of level-wise samples for the MLMC estimators with different L. For the standard (or single-level) MC estimator, the number of samples is decided according to (2.12) resulting in $N = \mathcal{O}(h^{-3})$. This means that the number of MC samples should be increased by a factor of eight with each grid refinement. The standard MC simulation was conducted on three grids: 16×16, 32×32 and 64×64 with samples 8, 64 and 512, respectively. The standard MC was not performed on the grid 128×128 due to prohibitively large computational cost, as we would need to compute about 4096 samples on this grid.

In Figure 7.7 (left), we show the mean relative errors in the expected value of v computed using the MC and MLMC estimators for Case 1. We observe that the plain MC estimator is slightly more accurate than the MLMC estimator for same finest grid h_L. The computational cost versus the accuracy for both methods is also shown in Figure 7.7 (right) and we observe that the MLMC estimator achieves the same accuracy for a lower computational cost compared to the MC estimator.

For reference, the predicted asymptotic cost of the MC and MLMC estimators for the considered α, β and γ are also presented. Similarly, the error and runtime from the two variance estimators are compared in Figure 7.8. Ideally, the cost of the MLMC estimator is expected to grow at half the rate of the MC estimator but this is not clearly visible for the multilevel estimator for the mean. This may very well be a pre-asymptotic effect. Nevertheless, the gains are more pronounced for the multilevel variance estimator and we clearly observe the cost scaling close to the predicted rate.

No. of levels $(L+1)$	Level-wise samples N_ℓ				
	N_0	N_1	N_2	N_3	N_4
1	8	-	-	-	-
2	64	8	-	-	-
3	512	64	8	-	-
4	4096	512	64	8	-
5 (ref)	32768	4096	512	64	8

Table 7.4: Number of samples used for the MLMC estimators with different L for the square duct flow. The 5-level MLMC estimator was utilized as the reference solution.
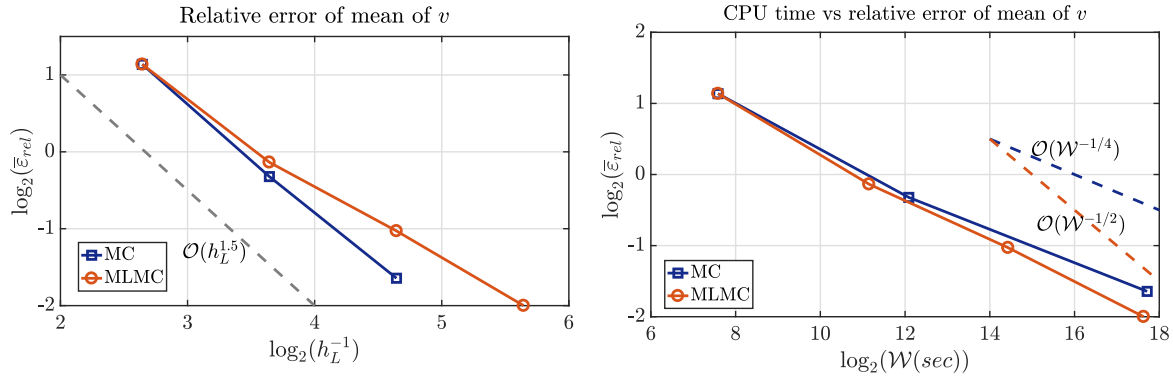
Figure 7.7: (Left) Comparison of the mean relative error $\bar{\varepsilon}_{rel}$ in the expected value of v for different meshes for Case 1. (Right) Computational work versus accuracy for the MC and MLMC estimators. Dotted lines show the predicted asymptotic cost for the MC (blue) and MLMC (red) estimators.

We now compare the stochastic solutions (mean and variance) for the RRST model computed using the 5-level estimator for Case 1 and 2 with the DNS data. The goal in this setting is to establish that the stochastic model is sufficiently general to (reliably) envelope DNS data at high probability as required for a good prior.

For the MLMC estimator, an appropriate spatial interpolation method is required to combine all expectations from the telescopic sum. To interpolate scalar fields from grid $\mathcal{D}_{\ell-1}$ to $\mathcal{D}_\ell$, a second-order spatial interpolation is employed. For instance, when using the multilevel estimator to compute $\mathcal{E}_L^{ML}[v_L]$, we proceed as follows. We begin by computing $\mathcal{E}_{N_0}^{MC}[v_0]$ on the coarsest grid $\mathcal{D}_0$. This is then interpolated to the next finer grid $\mathcal{D}_1$ and is added to the correction term $\mathcal{E}_{N_1}^{MC}[v_1 - v_0]$ resulting in a two-level estimate (a scalar field) $\mathcal{E}_1^{ML}[v_1]$. Similarly, this scalar field is further interpolated to the next grid and summed with the next correction term $\mathcal{E}_{N_2}^{MC}[v_2 - v_1]$. This process is repeated until the finest level is reached. Another possibility is to interpolate all expectations to the finest level and then add them together. Based on our experience, this may lead to interpolation artifacts in the final outcome.

In Figure 7.9 streamlines and magnitude of the in-plane velocities from the two cases are compared with the DNS data. We have observed that the size and the number of vortices are
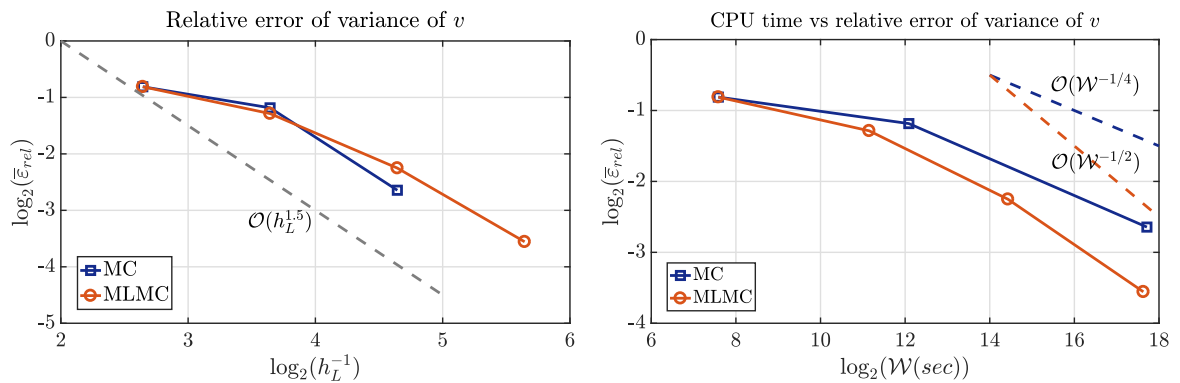
Figure 7.8: (Left) Comparison of the mean relative error $\bar{\varepsilon}_{rel}$ in the variance of v for different meshes for Case 1. (Right) Computational work versus accuracy for the MC and MLMC estimators.

sensitive to the correlation length; shorter lengths leading to more vortices. The secondary motions are entirely driven by the RRST model with magnitude of the velocities dependent on the value of the dispersion parameter. The mean $\pm$ standard deviation for the v velocity component at three locations is shown in Figure 7.10. We see that the two standard deviations envelopes the entire DNS velocity well. It is also pointed out that for Case 2, an even larger enveloping region is obtained. As mentioned earlier, we do not take into the account any available data and the hyper-parameters considered to generate the random Reynolds stresses were chosen arbitrarily. This high sensitivity of mean velocities with respect to change in Reynolds stresses is also demonstrated in [177] where an error of 1% in Reynolds stresses resulted in about 30% error in the mean velocity profile for the plain channel flow.

(a) Case 1   **(b)** Case 2   **(c)** DNS
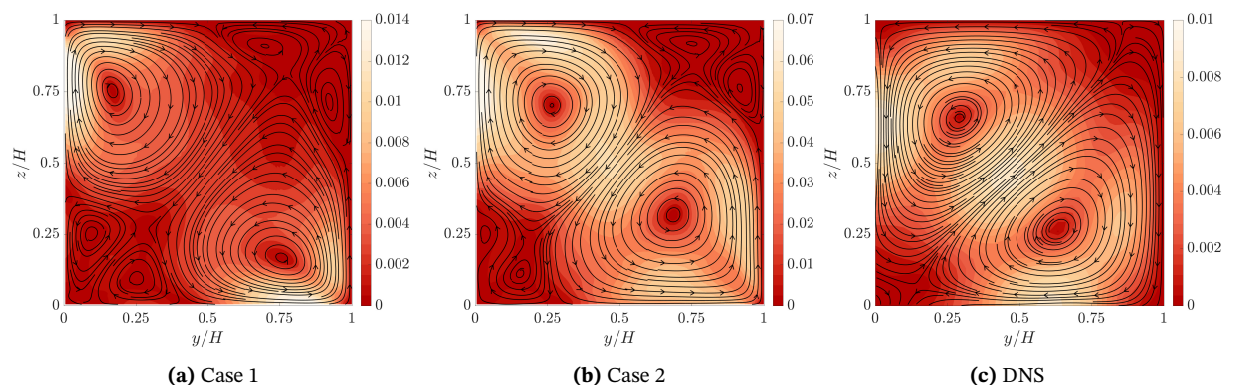
Figure 7.9: Comparison between the 5-level MLMC solution and the benchmark DNS data of the in-plane velocities v and w. Streamlines are constructed using $\mathcal{E}_L^{ML}[v_L]$ and $\mathcal{E}_L^{ML}[w_L]$ ($L = 4$) and the contour indicates the magnitude of the in-plane velocity vector ($\mathcal{E}_L^{ML}[v_L], \mathcal{E}_L^{ML}[w_L]$). Notice that with increase in dispersion δ an increase in the magnitude is observed.

We have propagated the uncertainty with about 10^5 degrees of freedom on the finest level due to six Reynolds stress components each sampled on a 128×128 grid. Note that there is a negligible change in the computational cost with an increase in dimensionality. Although, the uncertain dimension can be reduced by using the KL expansion one might still have to deal with a relatively large number of uncertainties rendering deterministic sampling methods such as the stochastic collocation method, generalized polynomial chaos, etc., impractical. However, multilevel variant of the Quasi-Monte Carlo (QMC) method, also based on deterministic sampling, can be highly efficient for this problem. QMC methods can in principle have dimension-independent convergence better than the plain MC, see [13, 33] for details.
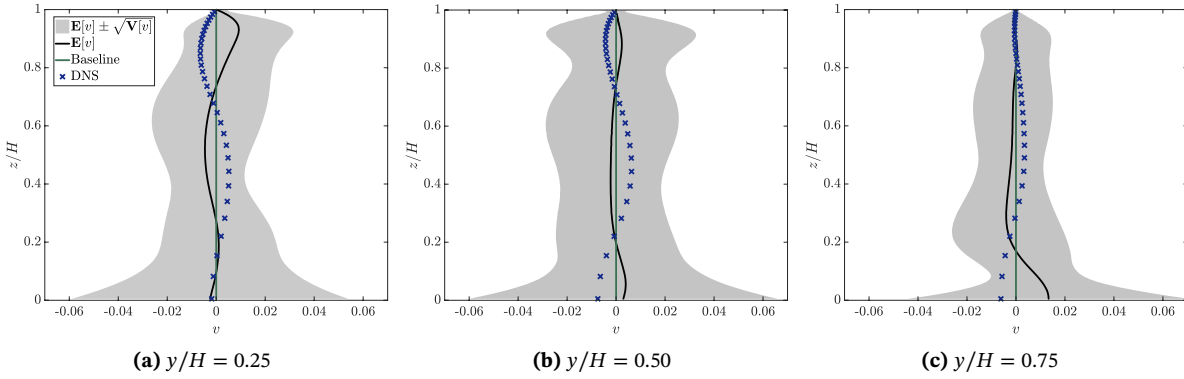
(a) $y/H = 0.25$ **(b)** $y/H = 0.50$ **(c)** $y/H = 0.75$

Figure 7.10: Mean $\mathcal{E}_L^{ML}[v_L]$ and variance $\mathcal{V}_L^{ML}[v_L]$ of the v-component of the velocity computed using the 5-level estimator for Case 1 at three spanwise locations $y/H = 0.25, 0.50$ and 0.75 with comparison to the baseline and DNS data.

7.4.2 Flow over periodic hills

The specification of the periodic hill geometry is adopted from [175]. The time-averaged flow from the DNS data is shown in Figure 7.11a. The size of the computational domain is $\mathcal{D}_x = 9H$ and $\mathcal{D}_y = 3.036H$ along the streamwise and wall-normal direction, respectively, with $H = 1$ denoting the hill height. The hill crest is situated at $(x/H, y/H) = (0, 1)$. Periodic boundary conditions are applied along the inlet and outlet boundaries and a solid stationary wall at the top and the bottom. The Reynolds number of the flow is given by $Re = u_b H/\nu = 2800$, where u_b is the average velocity above the hill crest and ν is the molecular viscosity. The numerical solutions are obtained on a curvilinear block-structured grid with two blocks of size $[0, 9] \times [0, 2]$ and $[0, 9] \times [2, 3.036]$, and refinement near the lower and upper walls.

Similar to the square duct case here also we use a pre-defined hierarchy of nested grids $\mathcal{D}_\ell$ such that we have $h_\ell \approx 0.5 h_{\ell-1}$. In Figure 7.11b, the two coarsest meshes are plotted. Also, the distribution of the y_ℓ^{+1} values (from the Launder-Sharma $k - \varepsilon$ model) along the lower wall for the five grids levels is depicted in Figure 7.12. All grid levels except the coarsest satisfy the criterion $y^{+1} < 1$. A separate grid hierarchy is used for the generation of the random fields. For a given grid level, these random fields are first sampled on a uniform rectangular mesh in a domain-sized $[0, 9] \times [0, 3]$ and are then interpolated to the cell-centers of the RANS simulation mesh. In Table 7.5, we list the specification for the different levels and the CPU times needed to obtain one sample on each level. For the considered combination of numerical schemes, we again observe a cost scaling roughly as $\mathcal{O}(h_\ell^{-3})$ or $\gamma = 3$.

Level (ℓ)	OF mesh	h_ℓ	RF mesh	CPU time (sec)
0	16×24	0.5625	24×8	0.26×10^2
1	32×48	0.2812	48×16	0.69×10^2
2	64×96	0.1406	96×32	6.82×10^2
3	128×192	0.0703	192×64	5.01×10^3
4	256×384	0.0352	384×128	4.70×10^4

Table 7.5: Specification of the MLMC grid hierarchy for the periodic hill case with $Re = 2800$. "OF mesh" denotes the simulation mesh in OpenFOAM and "RF mesh" denotes the grid used for the generation of the random eddy viscosity field. CPU time is the total time for one sample.
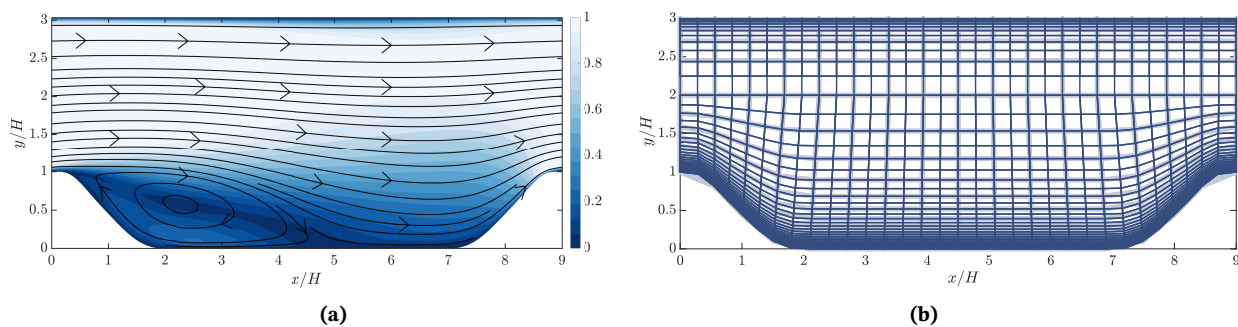
Figure 7.11: (a) Time-averaged flow on a periodic hill with Re = 2800 obtained from DNS data (Breuer et al. 2009) and (b) Nested curvilinear grids, the light blue lines depict 16x24 grid corresponding to $\ell = 0$ and dark blue lines depict 32x48 grid corresponding to $\ell = 1$.
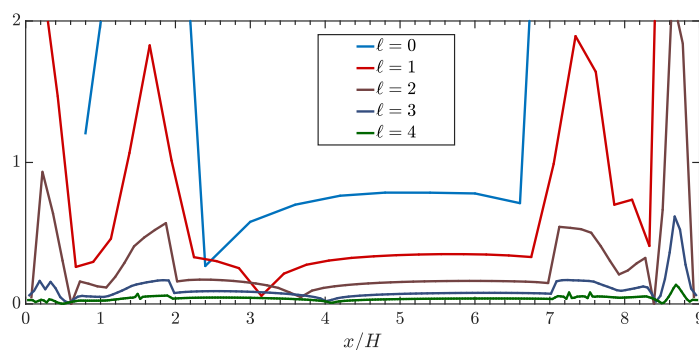

Figure 7.12: The y_ℓ^{+1} values along the lower wall computed from the baseline simulations.

MLMC with the REV model

We now analyze the performance of the MLMC method for the periodic hill flow using the random eddy viscosity (REV) stochastic model. To generate the samples of the random eddy viscosity two sets of parameters are utilized, denoted by Case 1 and Case 2, see Table 7.6. The two cases differ only in terms of the correlation length along the x- and y-directions. In Figure 7.13, we show an example of a REV field for each case along with the baseline field $\nu_t^{(bl)}(\mathbf{x})$ obtained from the converged solution from a $k - \epsilon$ model at the finest level $\mathcal{D}_4$ with 256×384 cells. Due to small correlation lengths for Case 2, we observe more peaks in the random eddy viscosity field with a relatively large magnitude. Here, for the sake of generality, we do not consider a periodic random eddy viscosity field, however, this could be easily implemented as the circulant embedding method naturally yields a periodic random field.

Parameter	Description	Case 1	Case 2
l_x/H	Correlation length along x-direction	1.5	0.6
l_y/H	Correlation length along y-direction	0.5	0.2
σ_c^2	Marginal variance of the random field	0.5	0.5

Table 7.6: Parameter sets to generate random eddy viscosity field for the periodic hill flow.

We perform a similar analysis as the square duct flow to obtain the MLMC parameters. We begin by comparing the FV error in the deterministic and stochastic versions of the problem for the streamwise velocity u and the wall shear stress τ^w in Figure 7.14. The error in the baseline
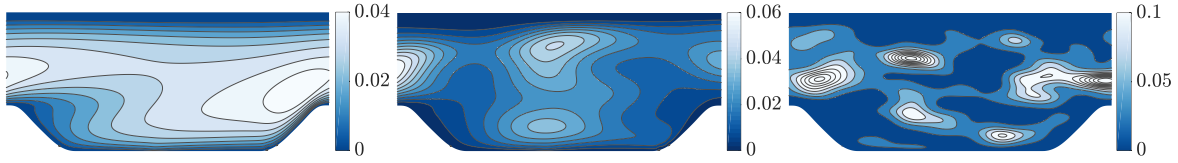
Figure 7.13: Baseline EV field from the Launder-Sharma $k - \epsilon$ model (left) and typical realizations of REV fields generated using the parameter sets from Case 1 (middle) and Case 2 (right).

converges as $\mathcal{O}(h_\ell)$ for both quantities of interest. The error in the random variables also decays at roughly the same rate. Here too, the slower convergence rate can be primarily attributed to complex curvilinear meshes. Also, note that the relative errors in Case 1 and 2 are very close, indicating that they result in similar mean solutions. The sampling variance on different levels is depicted in Figure 7.15. As expected the variance decays at a rate twice the discretization error coinciding with observations made in case of the square duct flow.
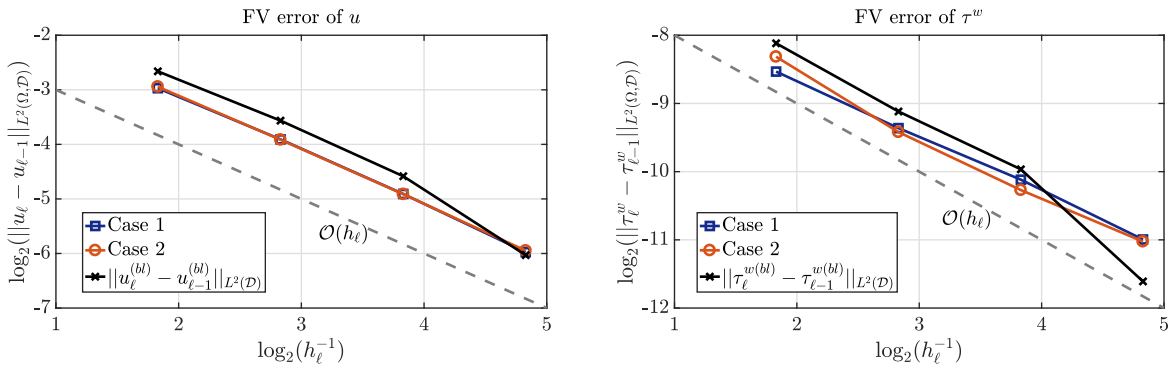
Figure 7.14: Convergence of the FV error with levels along with the error in the baseline solution. The dotted line denotes the empirical convergence rate of the baseline RANS simulations.
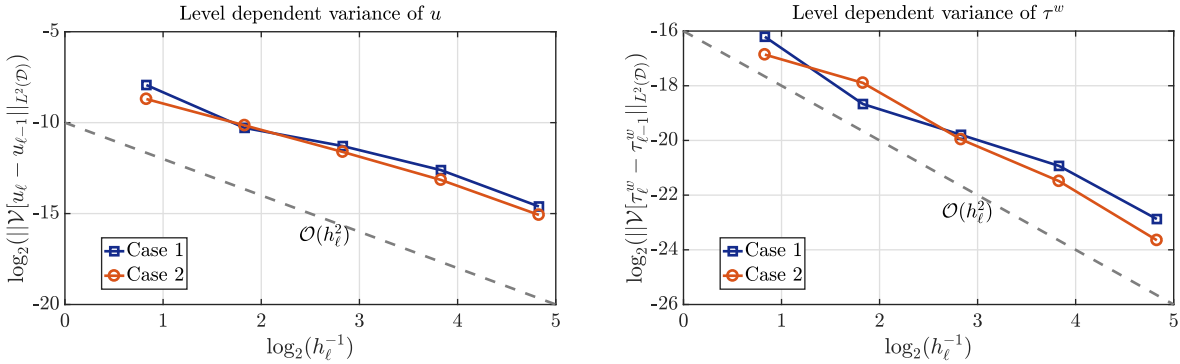
Figure 7.15: Convergence of the level-dependent variance for different grids. The dotted line depicts an $\mathcal{O}(h_\ell^2)$ convergence.

From the above study, we again illustrate that the rate from the baseline solution can provide a good estimate for the MLMC simulation parameters. Next, we analyze the relative errors in the MC and MLMC estimators for the streamwise velocity u in a similar fashion as for the square duct flow. As the quantity of interest we consider the streamwise velocity u and set $\alpha = 1$ and $\beta = 2\alpha = 2$ and $\gamma = 3$. Recall that, with these rates we end up in the third scenario $\gamma > \beta$ in (2.27), resulting in an asymptotically optimal MLMC estimator. The level-wise samples for the MLMC estimator are given by $N_\ell = N_L 2^{2(L-\ell)}$ with $N_L = 8$. The number of samples for the MLMC estimator with different L is given in Table 7.7. The reference solutions for the mean and variance $\mathcal{E}_{ref}[u]$ and

$\mathcal{V}_{ref}[u]$, respectively, are again based on the 5-level estimator. In case of the standard MC estimator, we follow $N = \mathcal{O}(h^{-2})$, thus the number of MC samples is increased by a factor of four with grid refinements. The standard MC simulation is conducted on four grids: 16×24, 32×48, 64×96 and 128×192 with samples 8, 32,128 and 512, respectively.

No. of levels ($L + 1$)	Level-wise samples N_ℓ				
	N_0	N_1	N_2	N_3	N_4
1	8	-	-	-	-
2	32	8	-	-	-
3	128	32	8	-	-
4	512	128	32	8	-
5 (ref)	2048	512	128	32	8

Table 7.7: Number of samples used for the MLMC estimators with different L for the flow over periodic hills. The 5-level MLMC estimator was utilized as the reference solution.

The mean relative error in the expectation of u approximated using the MC and MLMC methods is shown in Figure 7.16. The random eddy viscosity is based on Case 1. Both estimators are able to achieve similar accuracies, of order $\mathcal{O}(h_L)$. Also, the cost for both estimators scales similarly to the theoretical predictions. For $L = 3$, we see a speedup of up to 30 times using the MLMC estimator. In the case of the variance estimator in Figure 7.17, we observe slightly slower rates and the MLMC method appears to be a bit more accurate for the same grid. In terms of computational cost, similar gains are observed as for the expected value of u. We point out that for the MLMC estimator, the dominant cost comes from the finest level and as the number of samples N_L is a constant, we obtain a computational complexity of $\mathcal{O}(h_L^{-3})$. This is, up to a constant term, the same as solving one deterministic problem on the finest level, thus the MLMC estimator for this problem can be regarded as optimal.
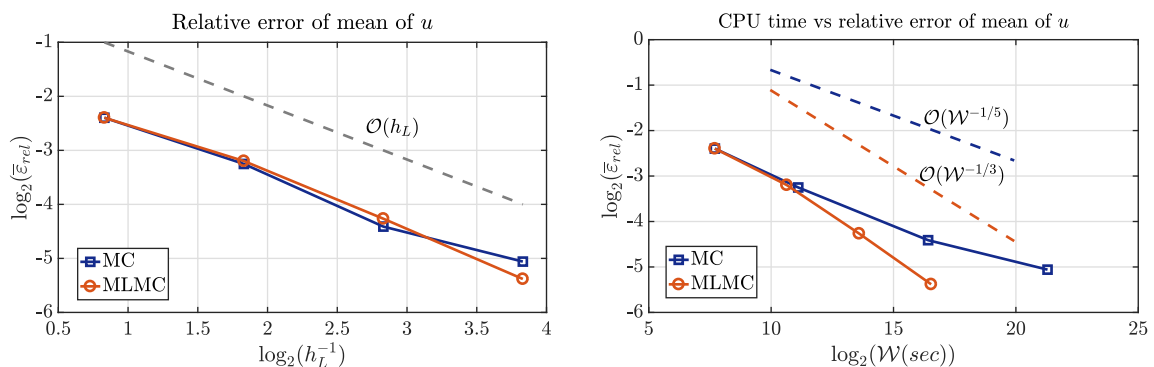

Figure 7.16: (Left) Comparison of the mean relative error $\bar{\varepsilon}_{rel}$ in the expected value of u for different meshes for Case 1. (Right) Computational work versus accuracy for the MC and MLMC estimators. Dotted lines show the predicted asymptotic cost for the MC (blue) and MLMC (red) estimators.

Next, we compare the reference solutions computed using the 5-level MLMC estimator with the REV model and DNS data. The main motivation of using the REV model was to obtain an uncertainty bound of the QoIs due to uncertainties arising from either the transport equations or the closure parameters. Therefore, we are interested in the computations of the variance field using the MLMC method. In Figure 7.18, the variance field for the streamwise velocity u for the two cases based on the 5-level estimator is shown. A relatively high variance is observed near boundary layers and near the recirculation zone around $0.5 < x/H < 4.5$. Case 1 is visibly able to generate a larger variance than Case 2 indicating larger length scales can produce larger variation. The mean $\pm$
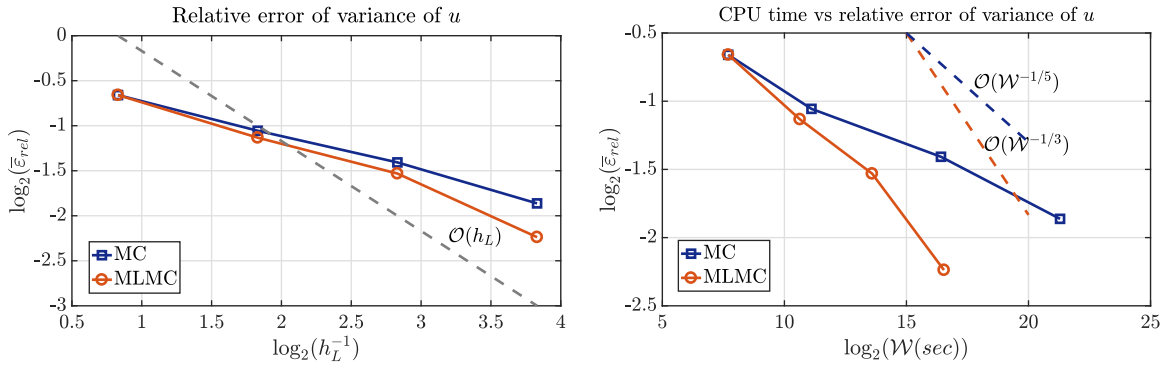
Figure 7.17: (Left) Comparison of the mean relative error $\bar{\varepsilon}_{rel}$ in the variance of u for different meshes for Case 1. (Right) Computational work versus accuracy for the MC and MLMC estimators.
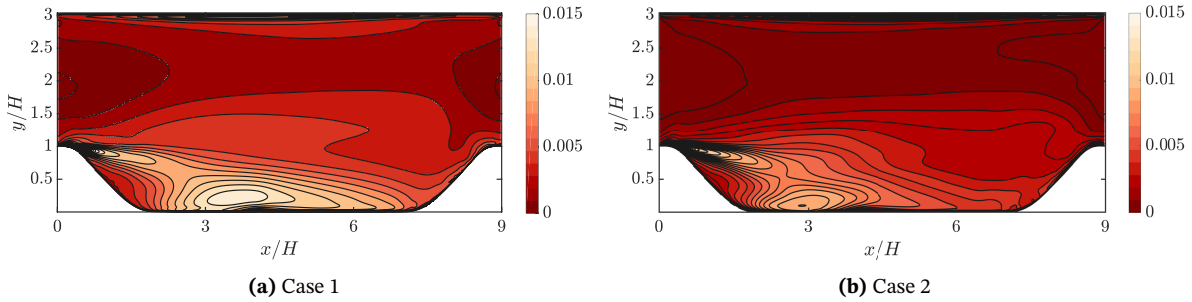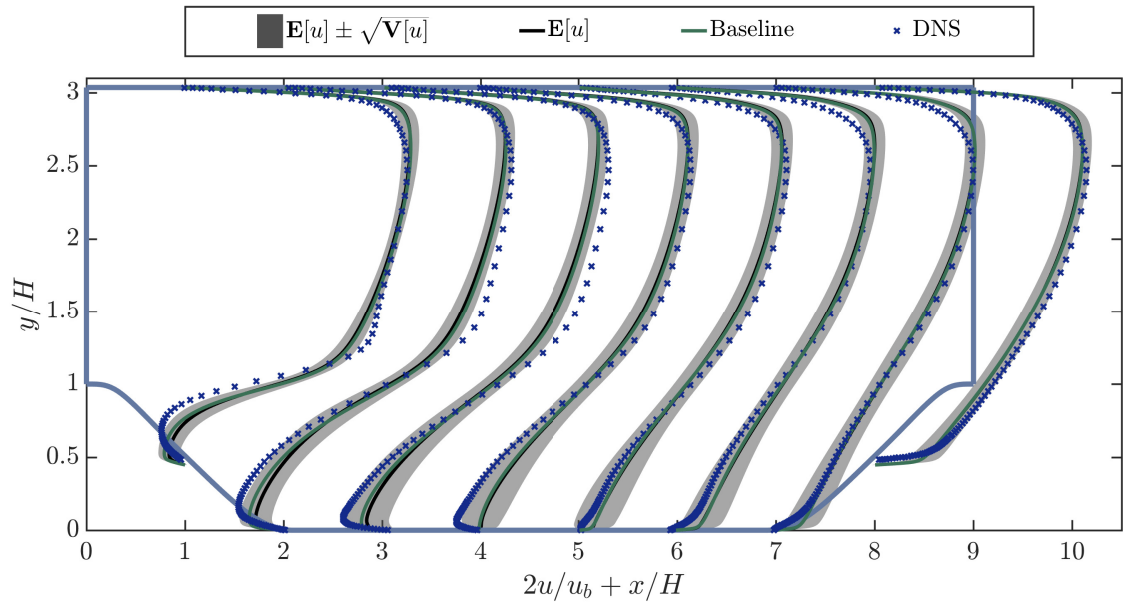

(a) Case 1

(b) Case 2

Figure 7.18: Variance field $\mathcal{V}_L^{ML}[u_L]$ for the streamwise velocity u computed using the 5-level estimator. Variance is large near top and bottom boundary layers.

standard deviation is compared with the DNS data at various locations in Figure 7.19. It can be seen that the MLMC velocity profiles are very close to the baseline RANS solution for both cases. Further, we observe that this stochastic model is less sensitive in the free shear layer as it fails to capture the DNS data very well. However, we have tested that a combination of larger marginal variance σ_c^2 and length scales $l_x/H, l_y/H$ can result in a larger uncertainty bound around the free shear layer. Despite randomly chosen turbulence models, interesting regions such as flow separation and reattachment can be detected from the variance field. Lastly, the mean and standard deviation obtained for the wall shear stress τ^w are also compared with the DNS data in Figure 7.20. Largest variances appear near the baseline reattachment point $x_{re}^{(bl)}$ near $x/H \approx 4$ for both the cases. For comparison, the DNS data is also plotted which falls within the $\pm$ one standard deviation bound of τ^w for both cases.

The reference solution presented above is based on 5×10^4 degrees of freedom because the random eddy viscosity field on the finest mesh was sampled on a 384×128 grid. Here too, the KL expansion based dimension reduction can be employed and may still result in a large number of random inputs, especially when the size of the domain is much larger than the correlation lengths.

MLMC with the RRST model

In the final numerical experiment, we test the performance of the MLMC method with the RRST model applied to the periodic hill test case. We use the same 5-grid hierarchy as was considered for the random eddy viscosity experiments to study the convergence of the bias and sampling error with respect to the levels. Also, the same number of OpenFOAM iterations was used to propagate the random stress tensor as was used to propagate the random eddy viscosity, thus, we have same the CPU time per sample as was given in Table 7.5 (neglecting the cost for sampling a single random

(a) Case 1
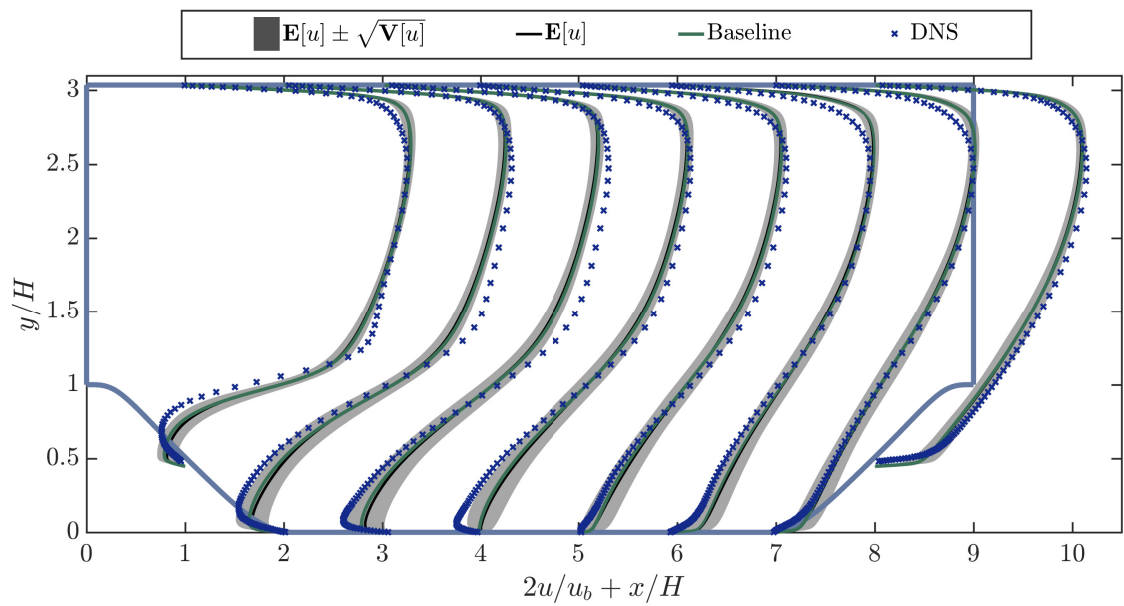
(b) Case 2

Figure 7.19: Mean and variance of the streamwise velocity computed using the 5-level estimator and comparison with DNS data at locations $x/H = 1, 2, 3, ..., 8$. Velocities are scaled by a factor of two to facilitate visualization.
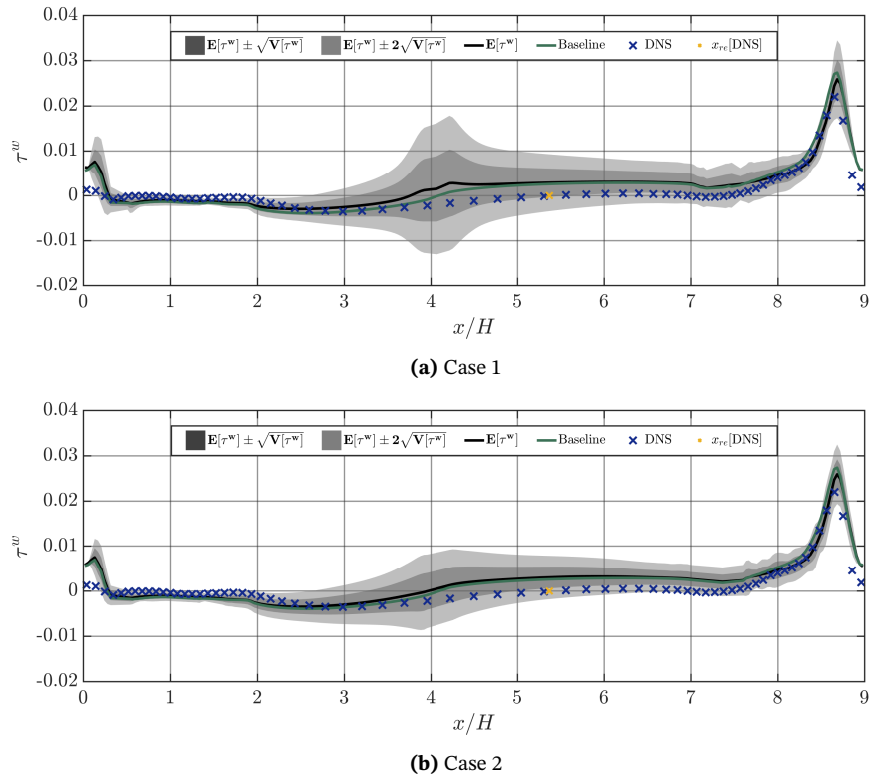
(a) Case 1

(b) Case 2

Figure 7.20: Mean and variance of the wall shear stress τ^w computed using the 5-level estimator and comparison with DNS data.

tensor field). The two parameter sets for generating the random tensor fields are listed in Table 7.8. For a fair comparison, we fix the blending parameter to $\xi = 0.6$ for both cases, although a higher blending is possible for the easier Case 1. Sample profiles of R_{12} for the two cases are compared in Figure 7.21 along with the baseline profile $R_{12}^{(bl)}$ (from the $k - \epsilon$ model). The effect of a larger dispersion and small correlation lengths is clearly visible for Case 2. We begin by analyzing the

Parameter	Description	Case 1	Case 2
l_x/H	Correlation length along x-direction	1.5	0.6
l_y/H	Correlation length along y-direction	0.5	0.2
σ_c^2	Variance of log-normal random field	1	1
$\delta(\mathbf{x})$	Dispersion parameter	0.2	0.4
N_{PC}	Order of polynomial chaos expansion	5	5
ξ	Blending factor	0.6	0.6

Table 7.8: Parameter sets to generate random Reynolds stress tensor for the flow over periodic hills.

convergence of the FV bias with grid refinements in Figure 7.22 for the streamwise velocity (left) and the wall shear stress (right). A first-order convergence is seen for the first four levels, similar to the REV model. However, for both cases the error is not reduced up to the discretization accuracy on the finest 256×384 grid. Similar behaviour is observed for the level-dependent variance in Figure 7.23, where the fifth level exhibits a larger variance compared to the fourth level. As we have similar rates as for the α, β, γ as the REV model, we use the same number of MLMC samples, from Table 7.7. Similarly, for the plain MC method 8, 32, 128 samples are used for the $16 \times 24, 32 \times 48$ and
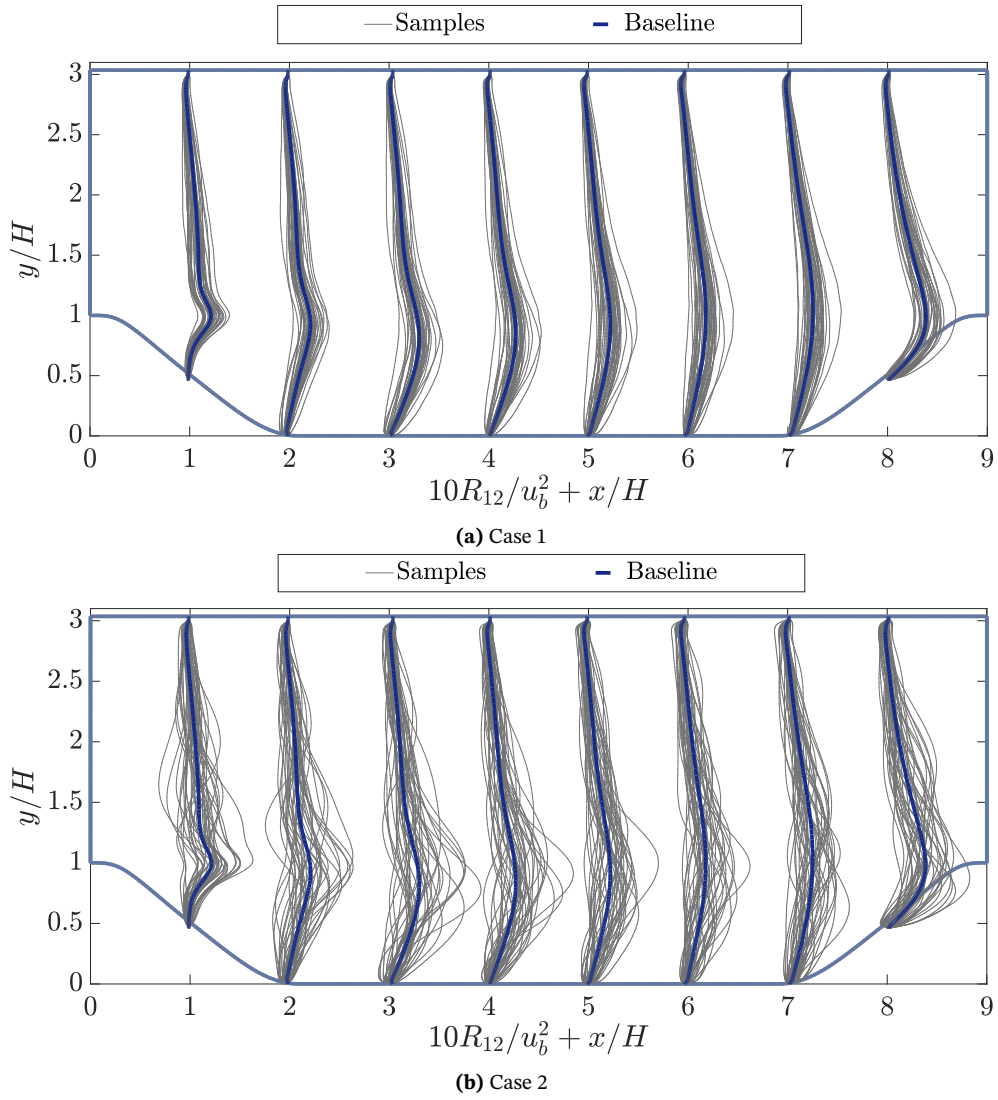
(a) Case 1

(b) Case 2

Figure 7.21: Comparison of sample profiles of R_{12} at different locations along with baseline values.
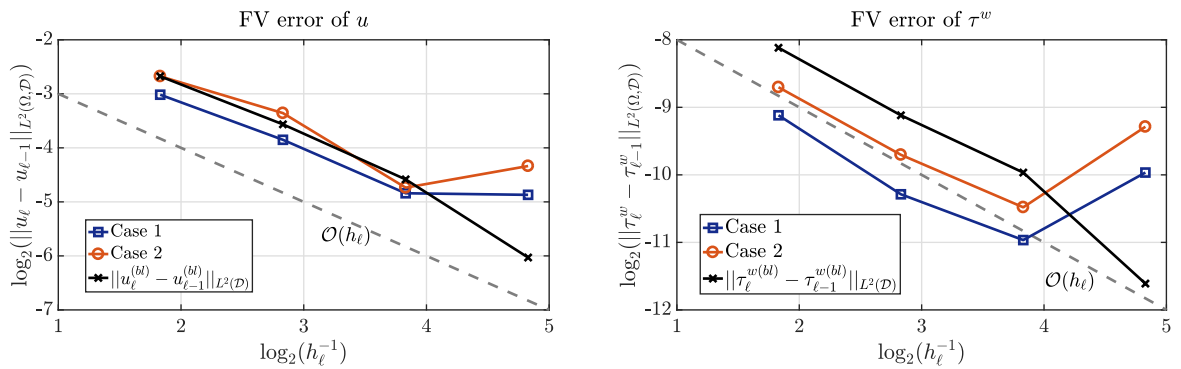

Figure 7.22: Convergence of the FV error with levels for the RRST model along with the error in baseline solution. The dotted line depicts $\mathcal{O}(h_\ell)$ convergence.
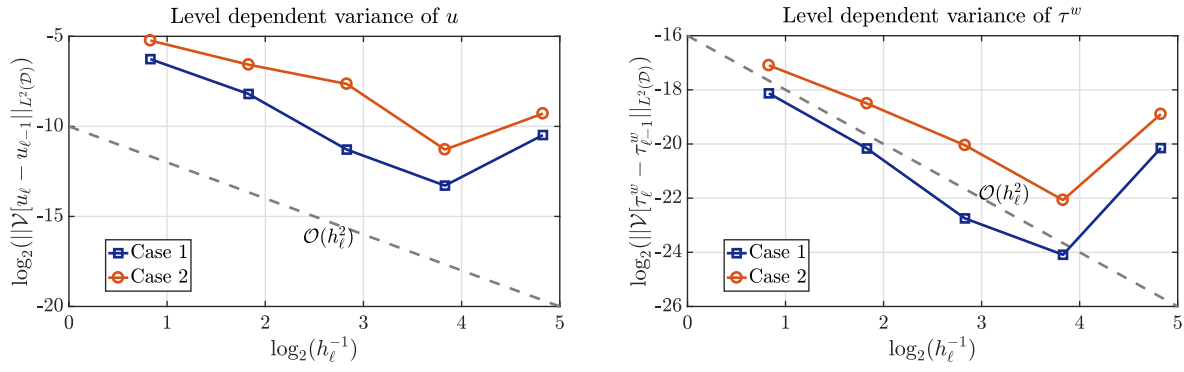
Figure 7.23: Convergence of the level-dependent variance for the RRST model. The dotted line depicts an $\mathcal{O}(h_\ell^2)$ convergence.
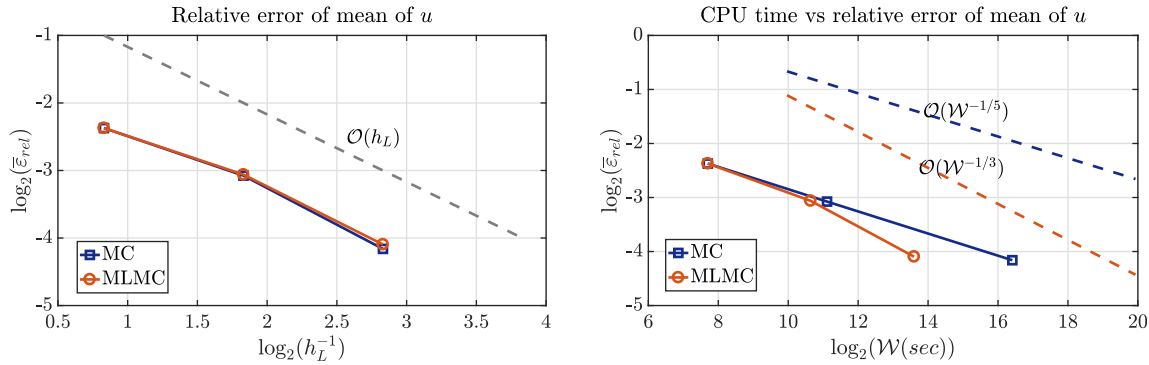

Figure 7.24: (Left) Comparison of the mean relative error $\bar{\varepsilon}_{rel}$ in the expected value of u for different meshes for Case 1. (Right) Computational work versus accuracy for the MC and MLMC estimators. Dotted lines show the predicted asymptotic cost for the MC (blue) and MLMC (red) estimators.

64×96 grids, respectively. The reference solutions for the mean and variance, $\mathcal{E}_{ref}[u]$ and $\mathcal{V}_{ref}[u]$, are based on a 4-level Monte Carlo estimator as the fifth level does not provide any improvement in the accuracy (for the considered solver). In figures 7.24 and 7.25, we show the mean relative errors and cost scaling for the mean and variance for Case 1. The speed-up is similar to the REV model and close to the theoretically predicted rates. The variance fields computed using the 4-level Monte Carlo for the two cases are presented in Figure 7.26. Larger variances are observed at locations where the effect of the turbulence is high, for example, near boundary layers and around locations where the flow starts to separate. The mean $\pm$ standard deviation of u at different locations is compared with the baseline and DNS data in Figure 7.27. As expected, a larger enveloping region is obtained for larger dispersion δ. The mean $\pm$ two standard deviations for the wall shear stress is also plotted in Figure 7.28. Again, a high variation is observed near the reattachment points obtained from the RANS simulation. We see that the DNS data falls within 2 standard deviations for both cases. We remind readers that the standard deviations observed are underestimated as the random tensor only contributes 60% of the propagated Reynolds stress tensor. For both quantities of interest, the observed means are very close to the baseline RANS solution, possibly indicating approximately linear dependence of u on the randomized RST.

7.5. Conclusions

In this work, we undertook the first steps towards solving high-dimensional stochastic formulations of RANS turbulence models using the multilevel Monte Carlo method. We demonstrated the efficiency of the MLMC method using two stochastic models based on a perturbation of the baseline
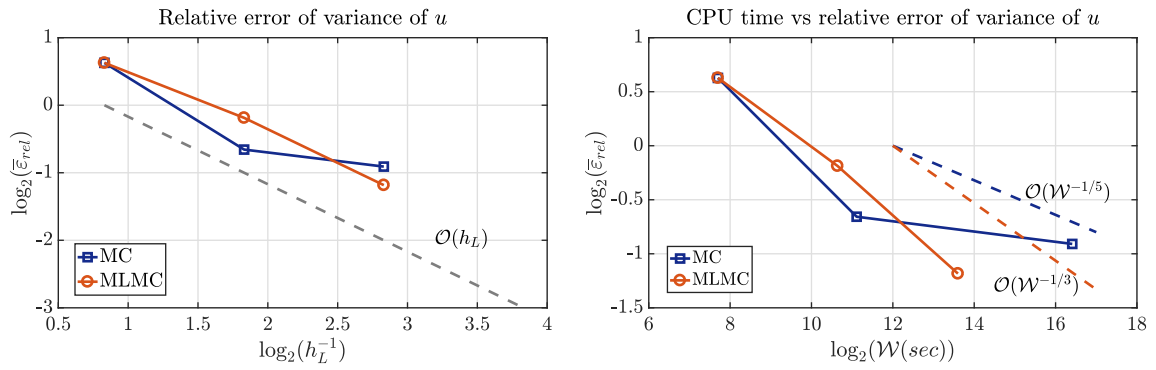
Figure 7.25: (Left) Comparison of the mean relative error $\bar{\varepsilon}_{rel}$ in the variance of u for different meshes for Case 1. (Right) Computational work versus accuracy for the MC and MLMC estimators.
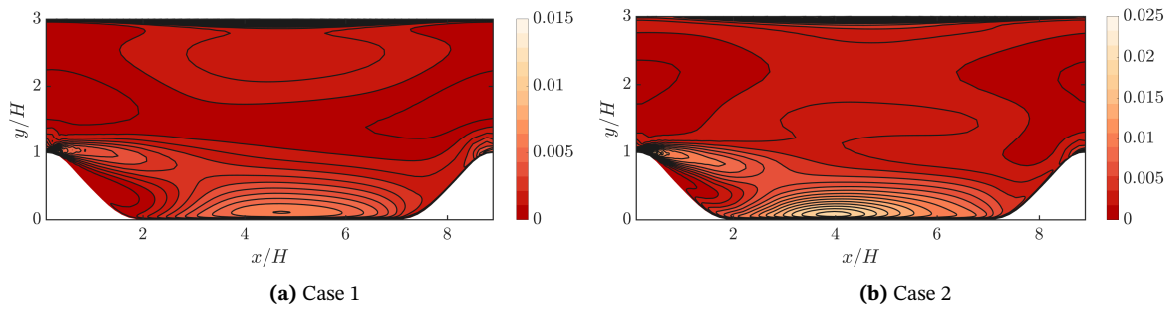

Figure 7.26: Variance field $\mathcal{V}_L^{ML}[u_L]$ for the streamwise velocity u computed using the 4-level estimator.

eddy viscosity field and the baseline Reynolds stress tensor field. The MLMC estimator proposed is only slightly more involved than the plain MC estimator but results in a large speed-up. The savings afforded by employing coarser levels allowed us to incorporate finer meshes thereby enabling the computation of the mean and variance with higher accuracy. We demonstrated that for QoIs for which the level-dependent variance decays slower than the growth rate of the computational cost with level (i.e. $\gamma > \beta$), an optimal MLMC estimator can be achieved. For two benchmarking problems, we utilized a nested and geometric hierarchy of grids. This is not a necessary criterion and a valid MLMC estimator can be constructed on any hierarchy that consists of levels with increasing cost and accuracy. A more sophisticated approach to obtain finer levels in the MLMC hierarchy can be based on adaptively refining the mesh in regions where a large numerical error is observed. We would like to point out that there is a negligible difference in terms of the computational cost between, the REV and RRST models, but the implementation of the latter model is more involved. Especially, obtaining a robust solver with respect to random Reynolds stress tensors is challenging. The continuation solver utilized in this chapter is moderately successful but suffers from convergence issues on very fine grids as well as when the random tensors are sampled from high-variance parameter sets.

This chapter presented the MLMC method as an efficient uncertainty propagation tool without taking into account any available data. A natural extension would be the development of multilevel variants of the Markov Chain Monte Carlo (MCMC) method to obtain a data-informed prediction [178]. For such algorithms, the random matrix approach can act as a better prior than the random eddy viscosity model as it circumvents the Boussinesq approximation. Currently, to the authors' knowledge, no data-based approach exists that takes into account the uncertainty in the full Reynolds tensor field.
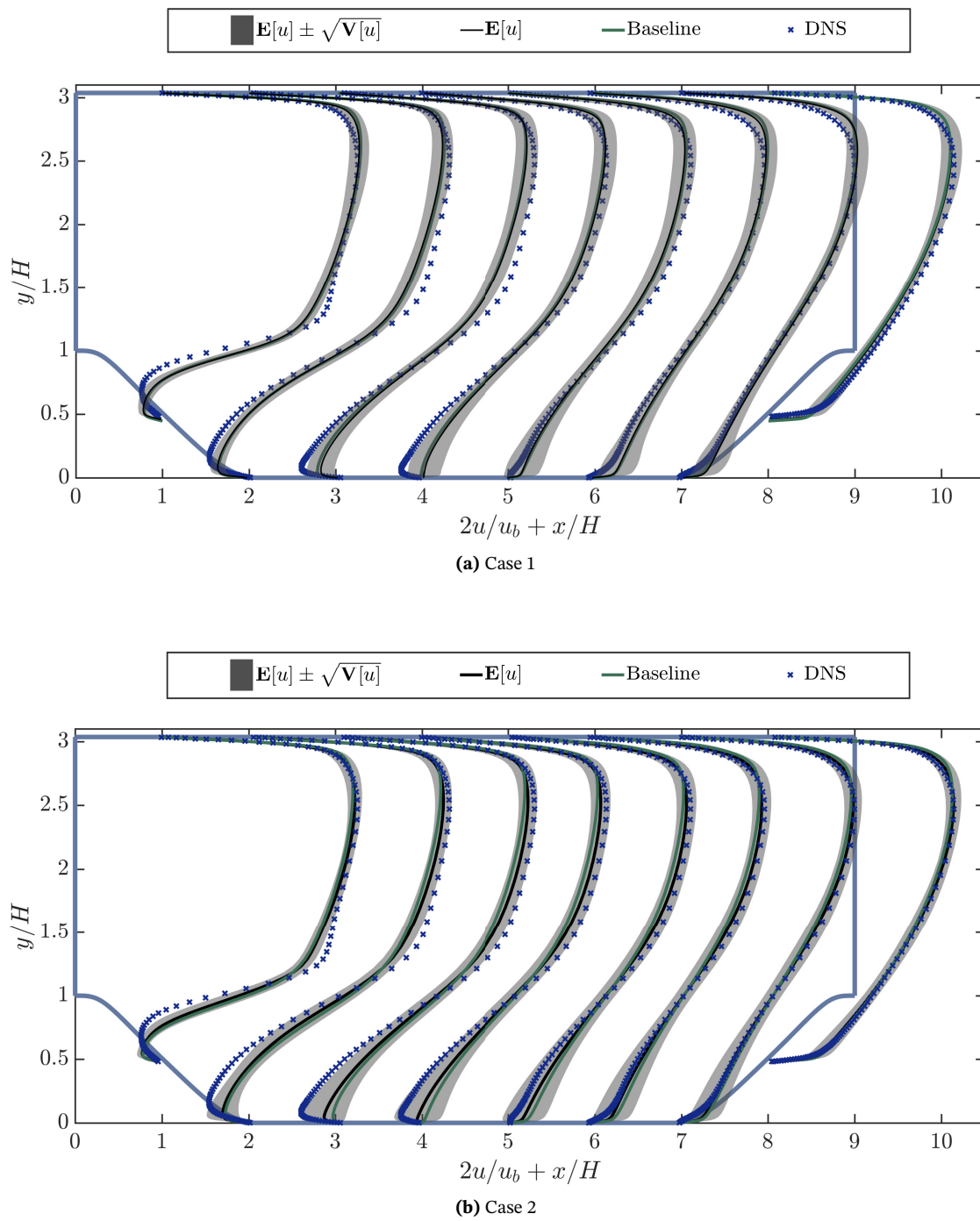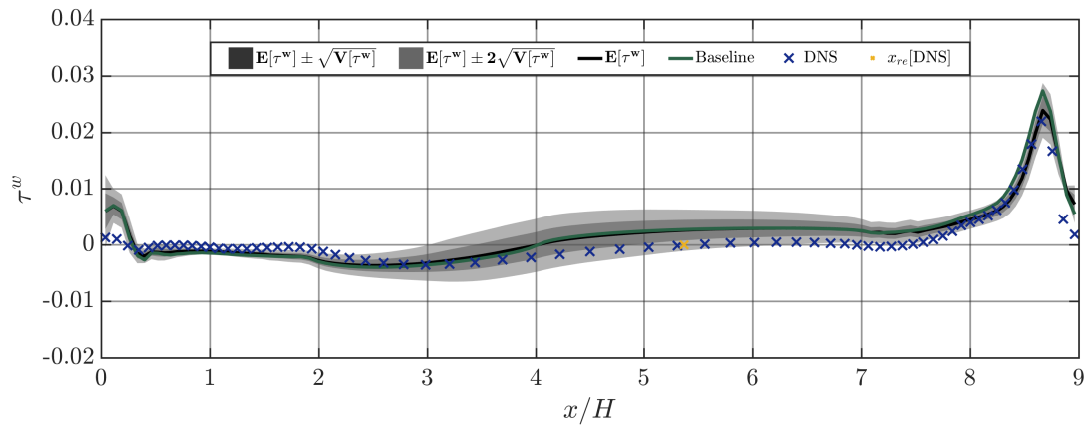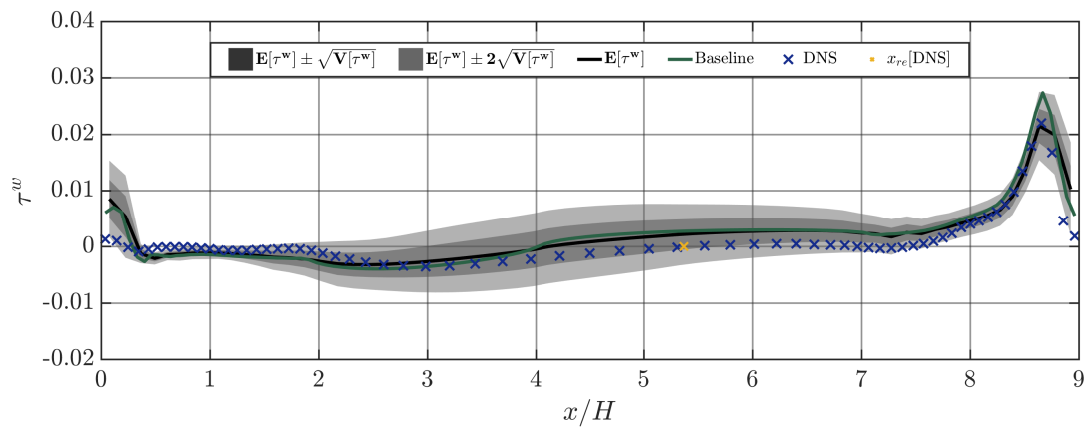
(a) Case 1

(b) Case 2

Figure 7.27: Mean and variance of the streamwise velocity computed using the 4-level estimator and comparison with DNS data at locations $x/H = 1, 2, 3, ..., 8$. Velocities are scaled by a factor of two to facilitate visualization.

(a) Case 1

(b) Case 2

Figure 7.28: Mean and variance of the wall shear stress τ^w computed using the 4-level estimator and comparison with DNS data.

Appendix

7.A. Projection of Reynolds Stresses on a Barycentric triangle

Reynolds stresses can be divided into an isotropic part $\frac{2}{3}k\delta_{ij}$ and an normalized anisotropic component given by

$$A_{ij} := \frac{R_{ij}}{2k} - \frac{\delta_{ij}}{3}, \tag{7.29}$$

$$\Rightarrow A_{ij} \in \begin{cases} [-1/3, 2/3] \text{ for } i = j, \\ [-1/2, 1/2] \text{ for } i \neq j, \end{cases} \tag{7.30}$$

forming the entries of a symmetric and deviatoric anisotropy tensor $\mathbf{A}$. Utilizing the eigenvalue decomposition, the anisotropy tensor $\mathbf{A}$, can be expressed as

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \tag{7.31}$$

where $\mathbf{V} = [v_1, v_2, v_3]$ with three mutually orthonormal eigenvectors v_i and the corresponding eigenvalue matrix $\mathbf{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \lambda_3]$ with $\lambda_1 + \lambda_2 + \lambda_3 = 0$ and ordering such that $\lambda_1 > \lambda_2 > \lambda_3$.

In physical terms, quantities $k, \mathbf{V}$ and $\mathbf{\Lambda}$ represent the magnitude, shape and orientation of the Reynolds stress, respectively. The state of the turbulence anisotropy can be visualized using a barycentric triangle [176]. This requires mapping the eigenvalues to the barycentric coordinates, C_{1c}, C_{2c}, C_{3c}, using linear relations:

$$C_{1c} = \lambda_1 - \lambda_2, \qquad C_{2c} = 2(\lambda_2 - \lambda_3), \qquad C_{3c} = 3\lambda_3 + 1, \quad \Rightarrow C_{1c} + C_{2c} + C_{3c} = 1. \tag{7.32}$$

Reynolds stress anisotropy is said to attain a limiting state when one of these components equals 1. Therefore, $C_{1c} = 1$ represents 1-component turbulence, $C_{2c} = 1$ represents 2-component turbulence and $C_{3c} = 1$ represents 3-component turbulence. One can express the anisotropy states in Cartesian coordinates using a barycentric triangle with the vertices $(x_{1c}, y_{1c}), (x_{2c}, y_{2c})$ and (x_{3c}, y_{3c}), corresponding to the three limiting states. Now, any anisotropy tensor can be projected into barycentric triangle via the convex combination of the three limiting states:

$$x = x_{1c}C_{1c} + x_{2c}C_{2c} + x_{3c}C_{3c}, \tag{7.33}$$
$$y = y_{1c}C_{1c} + y_{2c}C_{2c} + y_{3c}C_{3c}. \tag{7.34}$$

This transformation enables us to analyse the states of the Reynolds stresses generated using the random matrix approach. These perturbed Reynolds stresses should lie on, or within, this triangle to be physically realizable. The contours in Figure 7.4 are generated by making bins of equal size inside the barycentric triangle and plotting the normalized frequency for each bin.

8
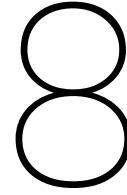
Conclusion and outlook

In this thesis, we study multilevel UQ techniques to efficiently solve a number of stochastic fluid flow problems. The priority has been on establishing efficient numerical schemes for the propagation of uncertainties. We have shown that for some problems, the multilevel approach is a highly feasible way to compute stochastic moments as the cost can be brought down to a small constant factor of one deterministic solve. Next, we draw some conclusions based on each research question addressed in this work.

8.1. General conclusion

8.1.1 High-order schemes for MLMC

The MLMC estimator is most expensive when the decay rate of level-dependent variance is slower than the growth rate of the computational cost per sample, i.e. $\beta < \gamma$. High-order discretization schemes can increase the value of β and we may end up with $\beta > \gamma$, resulting in the best possible MLMC complexity of $\mathcal{O}(\varepsilon^{-2})$ where ε is the specified tolerance of the MLMC estimator. However, for many problems, a high-order discretization scheme may result in a rather dense linear system which is comparatively more expensive to solve than the lower-order schemes (on the same grid) resulting in a larger γ. Due to this, we may not get any improvement in the asymptotic cost of the MLMC estimator. Therefore, it is also important to develop a fast solution method for high-order schemes in the MLMC setting. In Chapter 3, we investigated this research question in detail for the Darcy flow problem. We used a fourth-order discretization scheme instead of the conventionally used second-order scheme resulting in an increased value of β. For solving the linear system arising from the fourth-order discretization, we employed a defect correction strategy to a multigrid solver based on second-order discretization. As the defect correction step is inexpensive, we managed to keep the value of γ for fourth-order discretization the same as for the second-order; leading to a significant speedup in the MLMC simulations [23].

8.1.2 Local Fourier Analysis for Stochastic PDEs

The asymptotic convergence factors of multigrid solvers can be accurately predicted using LFA. Conventional LFA techniques are typically based on constant coefficient discretization, whereas for stochastic PDEs, we encounter variable coefficient fields due to randomness. In Chapter 4, we proposed a generalized LFA technique that is suitable for PDEs with variable and jumping coefficients. The generalization was achieved by employing new basis functions from the Fourier

space rather than using the standard Fourier modes. The prediction capability of the proposed LFA technique was demonstrated for a variety of challenging jumping coefficient problems as well as for highly heterogeneous random fields. We also showed that the aforementioned LFA method can be employed to compute the expected computational cost of the MLMC simulation [24].

8.1.3   MLMC for transport in the Darcy-Stokes flow

A multigrid MLMC method was described for solving the stochastic transport in a coupled Darcy-Stokes system in Chapter 5. Based on experience from previous chapters, we developed an efficient monolithic multigrid solver for the coupled Darcy-Stokes flow with heterogeneous permeability fields. The key ingredient of this solver was the Uzawa smoother that employs an equation-wise decoupled relaxation for the pressure and velocity unknowns. For the velocity, a symmetric Gauss-Seidel iteration is employed whereas, for the pressure, a Richardson iteration is applied. The choice of relaxation parameter for the Richardson iteration took into account the local fluctuations in the permeability field. Using numerical experiments, we showed the proposed monolithic multigrid solver is able to achieve the same efficiency for highly heterogeneous permeability fields as observed for the constant permeability case. A fast time-stepping of the advection-dominated transport equation was achieved using an Alternating Direction Implicit solver. This solver is highly suited for the MLMC method as it is a second-order accurate implicit scheme which allows us to incorporate very coarse temporal grids in the MLMC hierarchy [25].

8.1.4   MLMC for variably saturated flows

A stochastic extension for variably saturated flows was studied in Chapter 6. The stochastic formulation was derived by modeling the soil parameters as spatially correlated non-Gaussian random fields. A combination of the modified Picard iteration with a cell-centered multigrid was developed to solve the discrete Richards' equation. The performance of this solver was studied for a range of values of soil parameters and a cost map was derived for the considered range. For computing the statistical moments of the solution of Richards' equation, a parametric continuation multilevel Monte Carlo estimator was proposed. This estimator is highly practical for this problem, as it requires solving the strongly nonlinear problem target problems only on the finest level, where the solver is robust and solves simpler nonlinear problems on the coarse grid levels for a variance reduction [26].

8.1.5   MLMC for stochastic turbulence modeling

The MLMC method was utilized for solving high-dimensional stochastic formulations of RANS turbulence models in Chapter 7. Two stochastic RANS models based on a local perturbation of the baseline eddy viscosity field and the baseline Reynolds stress tensor field were considered. We showed that the random eddy viscosity model is suitable for quantifying uncertainties due to unknown closure coefficients for the turbulence models. The random Reynolds stress tensor model is more general and is applicable to flows where the assumption of linear stress-strain relation is insufficient. These models were tested on a selection of two relevant geometries and the uncertainty interval obtained was compared with the reference LES and DNS data. It was observed that the two models are sufficiently general and can reliably bound the possible flow behavior [27].

8.2. Extensions, outlook

There are a number of open research questions that were not addressed in this work. In this section, we outline some of these questions and recommend possible extensions of the work done in this thesis.

8.2.1   LFA for different types of PDEs and discretization schemes

In Chapter 4, we restricted ourselves to a basic cell-centered multigrid for elliptic PDEs to demonstrate the accuracy of the predictions of the novel LFA technique. We emphasize that this approach can also be used for a wider range of problems, discretizations, and multigrid methods. The proposed LFA technique allows us to deal with several challenging problems for which it is not easily possible to apply the classical LFA. For instance, this approach can be extended to higher-order cell-centered and vertex-centered discretizations and corresponding multigrid methods, or even to multigrid methods for staggered grid discretizations of coupled PDE systems, such as the Darcy-Stokes problem, see [25, 98]. Moreover, it should also be possible to apply it to triangular grids and finite element discretizations [179, 180].

8.2.2   Multilevel solvers for coupled PDEs

Using the insights from the monolithic multigrid solver developed for the stochastic Darcy-Stokes flow, we can develop an efficient monolithic solver for the coupled Stokes flow and deformable porous medium with highly heterogeneous permeability. A monolithic multigrid for this problem was already proposed in [181] but with constant permeability field. Our approach of using local fluctuations in the permeability to obtain optimal relaxation parameters can prove to be very effective in this case as well. Another interesting system that should be studied in the stochastic setting is the coupling of the Richards' equation with the Stokes flow. An efficient solver for this problem can be derived by combining the modified Picard iteration with the monolithic multigrid solver for the Darcy-Stokes flow.

8.2.3   Theoretical foundations of parametric continuation MLMC method

There are a number of research questions that need to be investigated in detail for the parametric continuation MLMC estimator. The first question concerns the perturbation strategy for the nonlinearity dictating parameters. In Chapter 6, we used a pre-defined hierarchy of parameters to compute the number of MLMC samples. This may not result in the maximum possible efficiency of the parametric continuation approach when applied in the MLMC setting. A more general approach needs to be formulated in order to find the optimal number of MLMC samples along with the parameter set for which the total cost of the MLMC estimator is minimum.

A second research question concerns the applicability of parametric continuation to other nonlinear problems that may exhibit bifurcation - where a small change in the perturbation of any parameter leads to a large difference in the solution profiles (for example, the Bratu problem). For such problems, the application of the parametric continuation MLMC method is not straightforward as the random samples on two levels $p_{\Theta_\ell, h_\ell}(\omega)$ and $p_{\Theta_{\ell-1}, h_{\ell-1}}(\omega)$ may not be correlated at all, therefore, it may not result in variance reduction. To deal with such issues, the arclength based continuation method [131] must be explored in a stochastic setting.

8.2.4   Bayesian calibration variants of MLMC for turbulence modeling

We showed that the two stochastic turbulence models in Chapter 7 were quite general and can serve as an accurate Bayesian prior for calibration and data-assimilation involving turbulence models. The next step would be the application of Bayesian calibration variants of MLMC methods, for example, Multilevel Markov-Chain Monte-Carlo (MLMCMC) [178], to this problem. For such algorithms, the random matrix approach [4] can act as a better prior than the random eddy viscosity model, as it circumvents the Boussinesq approximation.

All the previous methods were restricted by the accuracy of RANS computations: while the model error may be estimated, the absolute accuracy will not reach the level of LES "experiments".

Another important extension would be to develop techniques for uncertainty quantification of the LES model while limiting the number of LES computations to only a few or even to one. This may be achieved with a multi-fidelity approach, with RANS models on the coarse levels, using (for example) the control-variate approach of [182]. However, this would still require many LES computations. Another possibility would be to create an enriched coarse model using a RANS model and the Reynolds stress tensor information from a single LES model. The aim is to do this in such a way that the enriched RANS model at nominal conditions would reproduce the LES result (in the quantity of interest or the mean-flow).

All in all, there are plenty of open and interesting research questions left.

Bibliography

[1] M. C. Kennedy, A. O'Hagan, *Bayesian calibration of computer models*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63 (3) (2001) 425–464. `doi:10.1111/1467-9868.00294`.

[2] M. Emory, J. Larsson, G. Iaccarino, *Modeling of structural uncertainties in Reynolds-averaged Navier-Stokes closures*, Physics of Fluids 25 (11) (2013) 110 – 822. `doi:10.1063/1.4824659`.

[3] J. Ling, J. Templeton, *Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier-Stokes uncertainty*, Physics of Fluids 27 (8) (2015) 085–103. `doi:10.1063/1.4927765`.

[4] H. Xiao, J.-X. Wang, R. G. Ghanem, *A random matrix approach for quantifying model-form uncertainties in turbulence modeling*, Computer Methods in Applied Mechanics and Engineering 313 (Supplement C) (2017) 941 – 965. `doi:10.1016/j.cma.2016.10.025`.

[5] R. G. Ghanem, P. D. Spanos, *Stochastic finite elements: A spectral approach*, Springer, New York, 1991. `doi:10.1007/978-1-4612-3094-6`.

[6] I. Babuska, R. Tempone, G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM Journal on Numerical Analysis 42 (2) (2004) 800–825. `doi:10.1137/S0036142902418680`.

[7] D. Xiu, *Numerical Methods for stochastic computations: A spectral method approach*, Princeton University Press, Princeton, 2010.

[8] I. Babuška, F. Nobile, R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Journal on Numerical Analysis 45 (3) (2007) 1005–1034. `doi:10.1137/050645142`.

[9] S. Heinrich, *Multilevel Monte Carlo methods*, Large-Scale Scientific Computing, Third International Conference LSSC 2001, Lecture notes in Computer Science, 2170, Springer-Verlag, Sozopol, Bulgaria (2001) 58–67.

[10] M. B. Giles, *Multilevel Monte Carlo path simulation*, Operations Research 56 (3) (2008) 981–986. `doi:10.1287/opre.1070.0496`.

[11] F. Nobile, F. Tesei, *A multilevel Monte Carlo method with control variate for elliptic PDEs with log-normal coefficients*, Stochastic Partial Differential Equations: Analysis and Computations 3 (3) (2015) 398–444. `doi:10.1007/s40072-015-0055-9`.

[12] M. B. Giles, *Multilevel quasi-Monte Carlo path simulation*, Radon Series on Computational and Applied Mathematics 8 (2009) 1–18.

[13] F. Y. Kuo, Ch. Schwab, I. H. Sloan, *Multilevel quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients*, Foundations of Computational Mathematics 15 (2) (2015) 411–449. `doi:10.1007/s10208-014-9237-5`.

[14] U. Trottenberg, C. W. Oosterlee, A. Schuller, *Multigrid*, Elsevier Academic Press, San Diego, CA, 2000.

[15] A. Brandt, *Guide to multigrid development*, Multigrid Methods, Lecture notes in Mathematics 960 (eds W. Hackbusch and U. Trottenberg). Springer, Berlin (1982) 220–312.

[16] A. Brandt, O. E. Livne, *Multigrid techniques: 1984 guide with applications to fluid dynamics*, Vol. 67, SIAM, 2011.

[17] H. Darcy, *Les fontaines publiques de la ville de Dijon: exposition et application*, Victor Dalmont.

[18] J. P. Delhomme, *Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach*, Water Resources Research 15 (2) (1979) 269–280. `doi:10.1029/WR015i002p00269`.

[19] R. A. Freeze, *A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media*, Water Resources Research 11 (5) (1975) 725–741. `doi:10.1029/WR011i005p00725`.

[20] R. J. Hoeksema, P. K. Kitanidis, *Analysis of the spatial structure of properties of selected aquifers*, Water Resources Research 21 (4) (1985) 563–572. `doi:10.1029/WR021i004p00563`.

[21] L. Richards, *Capillary conduction of liquids through porous mediums*, Physics 1 (5) (1931) 318–333. `doi:10.1063/1.1745010`.

[22] W. N. Edeling, P. Cinnella, R. P. Dwight, *Predictive RANS simulations via Bayesian Model-Scenario Averaging*, Journal of Computational Physics 275 (Supplement C) (2014) 65 – 91. `doi:10.1016/j.jcp.2014.06.052`.

[23] P. Kumar, C. W. Oosterlee, R. P. Dwight, *A multigrid multilevel Monte Carlo method using high-order finite-volume scheme for lognormal diffusion problem*, International Journal for Uncertainty Quantification 7 (1) (2017) 57–81. `doi:10.1615/Int.J.UncertaintyQuantification.2016018677`.

[24] P. Kumar, C. Rodrigo, F. Gaspar, C. Oosterlee, On local fourier analysis of multigrid methods for pdes with jumping and random coefficients, SIAM Journal on Scientific Computing 41 (3) (2019) A1385–A1413. `doi:10.1137/18M1173769`.

[25] P. Kumar, P. Luo, F. J. Gaspar, C. W. Oosterlee, *A multigrid multilevel Monte Carlo method for transport in the Darcy-Stokes system*, Journal of Computational Physics 371 (2018) 382 – 408. `doi:10.1016/j.jcp.2018.05.046`.

[26] P. Kumar, C. Rodrigo, F. J. Gaspar, C. W. Oosterlee, *A parametric acceleration of multilevel Monte Carlo convergence for nonlinear variably saturated flows*, Preprint arXiv:1903.08741.

[27] P. Kumar, M. Schmelzer, R. P. Dwight, *Stochastic turbulence modeling in RANS simulations via multilevel Monte Carlo method*, Preprint arXiv:1811.00872.

[28] K. Cliffe, M. B. Giles, R. Scheichl, A. L. Teckentrup, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Computing and Visualization in Science 14 (2011) 3–15. `doi:10.1007/s00791-011-0160-x`.

[29] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numerica 24 (2015) 259–328. `doi:10.1017/S096249291500001X`.

[30] S. Mishra, Ch. Schwab, *Sparse tensor Multi-level Monte Carlo finite volume methods for hyperbolic conservation laws with random initial data*, Mathematics of Computation 81 (280) (2012) 1979–2018. `doi:10.1090/S0025-5718-2012-02574-9`.

[31] S. Mishra, Ch. Schwab, J. Šukys, *Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions*, Journal of Computational Physics 231 (8) (2012) 3365 – 3388. `doi:10.1016/j.jcp.2012.01.011`.

[32] S. Mishra, Ch. Schwab, J. Šukys, *Multi-level Monte Carlo finite volume methods for uncertainty quantification of acoustic wave propagation in random heterogeneous layered medium*, Journal of Computational Physics 312 (2016) 192–217. `doi:10.1016/j.jcp.2016.02.014`.

[33] F. Y. Kuo, Ch. Schwab, I. H. Sloan, E. Ullmann, *Multilevel quasi-Monte Carlo methods for lognormal diffusion problem*, Preprint arXiv:1507.01090v1.

[34] A.-L. Haji-Ali, F. Nobile, R. Tempone, *Multi-index Monte Carlo: when sparsity meets sampling*, Numerische Mathematik 132 (4) (2016) 767–806. `doi:10.1007/s00211-015-0734-5`.

[35] C.-H. Rhee, P. W. Glynn, *Unbiased estimation with square root convergence for SDE models*, Operations Research 63 (5) (2015) 1026–1043. `doi:10.1287/opre.2015.1404`.

[36] G. Detommaso, T. Dodwell, R. Scheichl, *Continuous level Monte Carlo and sample-adaptive model hierarchies*, SIAM/ASA Journal on Uncertainty Quantification 7 (1) (2019) 93–116. `doi:10.1137/18M1172259`.

[37] G. Sposito, W. A. Jury, V. K. Gupta, *Fundamental problems in the stochastic convection-dispersion model of solute transport in aquifers and field soils*, Water Resources Research 22 (1) (1986) 77–88. `doi:10.1029/WR022i001p00077`.

[38] I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl, I. H. Sloan, *Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications*, Journal of Computational Physics 230(10) (2011) 3668–3694. `doi:10.1016/j.jcp.2011.01.023`.

[39] J. Charrier, R. Scheichl, A. L. Teckentrup, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, SIAM Journal on Numerical Analysis 51 (1) (2013) 322–352. `doi:10.1137/110853054`.

[40] A. Barth, Ch. Schwab, N. Zollinger, *Multilevel Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numerische Mathematik 119 (1) (2011) 123–161. `doi:10.1007/s00211-011-0377-0`.

[41] A. L. Teckentrup, R. Scheichl, M. B. Giles, E. Ullmann, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numerische Mathematik 125 (3) (2013) 569–600. `doi:10.1007/s00211-013-0546-4`.

[42] M. B. Giles, *Improved multilevel Monte Carlo convergence using Milstein scheme*, Monte Carlo and Quasi-Monte Carlo methods 2006, Springer (2008) 343–358.

[43] M. Handcock, J. Wallis, *An approach to statistical spatial-temporal modeling of meteorological fields (with discussion)*, Journal of the American Statistical Association 89 (1994) 368–390. `doi:10.2307/2290832`.

[44] C. E. Rasmussen, *Gaussian processes for machine learning*, MIT Press, 2006.

[45] R. J. Adler, *The geometry of random fields*, Society for Industrial and Applied Mathematics, 2010. `doi:10.1137/1.9780898718980`.

[46] J. Galvis, M. Sarkis, *Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity*, SIAM Journal on Numerical Analysis 47 (5) (2009) 3624–3651. `doi:10.1137/080717924`.

[47] D. Gilbarg, N. S. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag Berlin Heidelberg. `doi:10.1007/978-3-642-61798-0`.

[48] J. Charrier, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM Journal on Numerical Analysis 50 (1) (2012) 216–246. `doi:10.1137/100800531`.

[49] G. Dagan, *Flow and transport in porous formations*, Springer, Berlin, Heidelberg, 1989. `doi:10.1007/978-3-642-75015-1`.

[50] P. Colella, M. R. Dorr, J. A. F. Hittinger, D. F. Martin, *High-Order, Finite-Volume Methods in Mapped Coordinates*, Journal of Computational Physics 230 (2011) 2952–2976. `doi:10.1016/j.jcp.2010.12.044`.

[51] P. McCorquodale, P. Colella, *A high-order finite-volume method for hyperbolic conservation laws on locally-refined grids*, Communications in Applied Mathematics and Computational Science 6 (1) (2011) 1–25.

[52] M. Vinokur, *An analysis of finite-difference and finite-volume formulations of conservation laws*, Journal of Computational Physics 81 (1) (1989) 1 – 52. `doi:10.1016/0021-9991(89)90063-6`.

[53] Q. Zhang, H. Johansen, P. Colella, *A fourth-order accurate finite-volume method with structured adaptive mesh refinement for solving the Advection-diffusion equation*, SIAM Journal on Scientific Computing 34 (2) (2012) B179–B201. `doi:10.1137/110820105`.

[54] M. Khalil, P. Wesseling, *Vertex-centered and cell-centered multigrid for interface problems*, Journal of Computational Physics 98 (1992) 1–10. `doi:10.1016/0021-9991(92)90168-X`.

[55] M. Mohr, R. Wienands, *Cell-centred multigrid revisited*, Computing and Visualization in Science 7 (2004) 129–140. `doi:10.1007/s00791-004-0137-0`.

[56] W. Auzinger, H. J. Stetter, *Defect corrections and multigrid iterations*, Multigrid Methods, Lecture notes in Mathematics 960, Springer, Berlin (1982) 327–351.

[57] P. W. Hemker, *Mixed defect correction iteration for the accurate solution of convection diffusion equation*, Multigrid Methods, Lecture notes in Mathematics 960, Springer, Berlin (1982) 485–501.

[58] J. Ruge, K. Stüben, *Efficient solution of finite difference and finite element equations by algebraic multigrid (AMG)*, D.J. Paddon, H. Holstein (Eds.), Multigrid Methods for Integral and Differential Equations, IMA Conference Series, Clarendon Press, Oxford (1985) 169–212.

[59] P. Robbe, D. Nuyens, S. Vandewalle, *Recycling Samples in the multigrid multilevel (Quasi-) Monte Carlo method*, Preprint arXiv:1806.05619.

[60] M. L. Ravalec, B. Noetinger, L. Y. Hu, *The FFT Moving Average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations*, Mathematical Geology 32 (6) (2000) 701–723. `doi:10.1023/A:1007542406333`.

[61] A. T. A. Wood, G. Chan, *Simulation of stationary Gaussian processes in* $[0,1]^d$, Journal of Computational and Graphical Statistics 3 (1994) 409–432. `doi:10.1080/10618600.1994.10474655`.

[62] C. R. Dietrich, G. N. Newsam, *Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix*, SIAM Journal on Scientific Computing 18 (1997) 1088–1107. `doi:10.1137/S1064827592240555`.

[63] J. Moulton, J. E. Dendy, J. M. Hyman, *The black box multigrid numerical homogenization algorithm*, Journal of Computational Physics 142 (1) (1998) 80 – 108. `doi:10.1006/jcph.1998.5911`.

[64] S. P. MacLachlan, J. D. Moulton, *Multilevel upscaling through variational coarsening*, Water Resources Research 42 (2). `doi:10.1029/2005WR003940`.

[65] B. Seynaeve, E. Rosseel, B. Nicola, S. Vandewalle, *Fourier mode analysis of multigrid methods for partial differential equations with random coefficients* , Journal of Computational Physics 224 (1) (2007) 132 – 149. `doi:10.1016/j.jcp.2006.12.011`.

[66] I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, Ch. Schwab, I. H. Sloan, *Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients*, Numerische Mathematik 131 (2) (2015) 329–368. `doi:10.1007/s00211-014-0689-y`.

[67] M. Bolten, H. Rittich, *Fourier analysis of periodic stencils in multigrid methods*, SIAM Journal on Scientific Computing 40 (3) (2018) A1642–A1668. `doi:10.1137/16M1073959`.

[68] J. Molenaar, *A simple cell-centered multigrid method for 3D interface problems*, Computers and Mathematics with Applications 31 (9) (1996) 25 – 33. `doi:10.1016/0898-1221(96)00039-9`.

[69] I. Yavneh, *Coarse-grid correction for nonelliptic and singular perturbation problems*, SIAM Journal on Scientific Computing 19 (5) (1998) 1682–1699. `doi:10.1137/S1064827596310998`.

[70] W. Hackbusch, *Multigrid Methods and Applications*, Springer, 1985.

[71] J. H. Bramble, R. E. Ewing, J. E. Pasciak, J. Shen, *The analysis of multigrid algorithms for cell centered finite difference methods*, Advances in Computational Mathematics 5 (1) (1996) 15–29. `doi:10.1007/BF02124733`.

[72] J. H. Bramble, J. E. Pasciak, J. Xu, *The analysis of multigrid algorithms with non-nested Spaces or non-inherited quadratic forms*, Mathematics of Computation 56 (193) (1991) 1–34. `doi:10.2307/2008527`.

[73] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, J. Ruge, *Adaptive Smoothed Aggregation (αSA) multigrid*, SIAM Review 47 (2) (2005) 317–346. `doi:10.1137/S1064827502418598`.

[74] K. Stüben, *Appendix A: An introduction to algebraic multigrid*, in Multigrid, U. Trottenberg, C. W. Oosterlee, and A. Schüller, Academic Press, San Diego, CA, 2001 (2001) 413–532.

[75] P. Vaněk, M. Brezina, J. Mandel, *Convergence of algebraic multigrid based on smoothed aggregation*, Numerische Mathematik 88 (3) (2001) 559–579. `doi:10.1007/s211-001-8015-y`.

[76] D. Braess, *Towards algebraic multigrid for elliptic problems of second order*, Computing 55 (4) (1995) 379–393. `doi:10.1007/BF02238488`.

[77] P. Vaněk, J. Mandel, M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing 56 (3) (1996) 179–196. `doi:10.1007/BF02238511`.

[78] R. E. Alcouffe, A. Brandt, J. E. Dendy, Jr., J. W. Painter, *The multigrid methods for the diffusion equation with strongly discontinuous coefficients*, SIAM Journal on Scientific and Statistical Computing 2 (4) (1981) 430–454. `doi:10.1137/0902035`.

[79] J. Dendy, *Black box multigrid*, Journal of Computational Physics 48 (3) (1982) 366 – 386. `doi:10.1016/0021-9991(82)90057-2`.

[80] J. Dendy, *Black box multigrid for non-symmetric problems*, Applied Mathematics and Computation 13 (3) (1983) 261 – 283. `doi:10.1016/0096-3003(83)90016-4`.

[81] P. Wesseling, *Cell-centered multigrid for interface problems*, Journal of Computational Physics 79 (1) (1988) 85 – 91. `doi:10.1016/0021-9991(88)90005-8`.

[82] J. E. Dendy, J. D. Moulton, *Black box multigrid with coarsening by a factor of three*, Numerical Linear Algebra with Applications 17 (2) 577–598. `doi:10.1002/nla.705`.

[83] S. Knapek, *Matrix-dependent multigrid homogenization for diffusion problems*, SIAM Journal on Scientific Computing 20 (2) (1998) 515–533. `doi:10.1137/S1064827596304848`.

[84] R. Wienands, C. W. Oosterlee, *On three-grid Fourier analysis for multigrid*, SIAM Journal on Scientific Computing 23 (2) (2001) 651–671. `doi:10.1137/S106482750037367X`.

[85] G. S. Beavers, D. D. Joseph, *Boundary conditions at a naturally permeable wall*, Journal of Fluid Mechanics 30 (01) (1967) 197–207. `doi:10.1017/S0022112067001375`.

[86] P. G. Saffman, *On the boundary condition at the surface of a porous medium*, Studies in Applied Mathematics 50 (2) (1971) 93–101. `doi:10.1002/sapm197150293`.

[87] P. Angot, *On the well-posed coupling between free fluid and porous viscous flows*, Applied Mathematics Letters 24 (6) (2011) 803–810. `doi:10.1016/j.aml.2010.07.008`.

[88] Y. Cao, M. Gunzburger, F. Hua, X. Wang, *Coupled Stokes-Darcy model with Beavers-Joseph interface boundary condition*, Communications in Mathematical Sciences 8 (1) (2010) 1–25.

[89] M. Discacciati, E. Miglio, A. Quarteroni, *Mathematical and numerical models for coupling surface and groundwater flows*, Applied Numerical Mathematics 43 (1-2) (2002) 57–74. `doi:10.1016/S0168-9274(02)00125-3`.

[90] H. P. Langtangen, K. A. Mardal, R. Winther, *Numerical methods for incompressible viscous flow*, Advances in Water Resources 25 (8) (2002) 1125–1146. `doi:10.1016/S0309-1708(02)00052-0`.

[91] W. J. Layton, F. Schieweck, I. Yotov, *Coupling fluid flow with porous media flow*, SIAM Journal on Numerical Analysis 40 (6) (2002) 2195–2218. `doi:10.1137/S0036142901392766`.

[92] A. Mikelic, W. Jäger, *On the interface boundary condition of Beavers, Joseph, and Saffman*, SIAM Journal on Applied Mathematics 60 (4) (2000) 1111–1127. `doi:10.1137/S003613999833678X`.

[93] B. Rivière, I. Yotov, *Locally conservative coupling of Stokes and Darcy flows*, SIAM Journal on Numerical Analysis 42 (5) (2005) 1959–1977. `doi:10.1137/S0036142903427640`.

[94] D. Vassilev, I. Yotov, *Coupling Stokes-Darcy Flow with Transport*, SIAM Journal on Scientific Computing 31 (5) (2009) 3661–3684. `doi:10.1137/080732146`.

[95] J. Maitre, F. Musy, P. Nigon, *A fast solver for the Stokes equations using multigrid with a Uzawa smoother*, in: Advances in multigrid methods, Springer, 1985, pp. 77–83. `doi:10.1007/978-3-663-14245-4_8`.

[96] R. Eymard, T. Gallouët, R. Herbin, *Finite volume methods*, Handbook of numerical analysis 7 (2000) 713–1018. `doi:10.1016/S1570-8659(00)07005-8`.

[97] F. J. Gaspar, Y. Notay, C. W. Oosterlee, C. Rodrigo, *A simple and efficient segregated smoother for the discrete Stokes equations*, SIAM Journal on Scientific Computing 36 (3) (2014) A1187–A1206. `doi:10.1137/130920630`.

[98] P. Luo, C. Rodrigo, F. J. Gaspar, C. W. Oosterlee, *Uzawa smoother in multigrid for the coupled porous medium and Stokes flow system*, SIAM Journal on Scientific Computing 39 (5) (2017) S633–S661. `doi:10.1137/16M1076514`.

[99] F. dell 'Isola, A. Madeo, P. Seppecher, *Boundary conditions at fluid-permeable interfaces in porous media: A variational approach*, International Journal of Solids and Structures 46 (17) (2009) 3150 – 3164. `doi:10.1016/j.ijsolstr.2009.04.008`.

[100] L. Payne, B. Straughan, *Analysis of the boundary condition at the interface between a viscous fluid and a porous medium and related modelling questions*, Journal de Mathématiques Pures et Appliquées 77 (4) (1998) 317 – 354. `doi:10.1016/S0021-7824(98)80102-5`.

[101] J. Bear, *Dynamics of fluids in porous media*, Courier Corporation, 2013.

[102] C. Zheng, P. P. Wang, *MT3DMS: A modular three-dimensional multi-species transport model for simulation of advection, dispersion, and chemical reactions of contaminants in ground-water systems*, Contract Report SERDP-99-1, US Army Engineer Research and Development, 1999.

[103] F. H. Harlow, J. E. Welch, *Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface*, The physics of fluids 8 (12) (1965) 2182–2189. `doi:10.1063/1.1761178`.

[104] B. P. Leonard, *A stable and accurate convective modelling procedure based on quadratic upstream interpolation*, Computer methods in applied mechanics and engineering 19 (1) (1979) 59–98. `doi:10.1016/0045-7825(79)90034-3`.

[105] B. Leonard, S. Mokhtari, *Beyond first-order upwinding: The ultra-sharp alternative for non-oscillatory steady-state simulation of convection*, International Journal for Numerical Methods in Engineering 30 (4) (1990) 729–766. `doi:10.1002/nme.1620300412`.

[106] B. Leonard, *The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection*, Computer methods in applied mechanics and engineering 88 (1) (1991) 17–74. `doi:10.1016/0045-7825(91)90232-U`.

[107] P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM Journal on Numerical Analysis 21 (5) (1984) 995–1011. `doi:10.1137/0721062`.

[108] B. Koren, *A robust upwind discretization method for advection, diffusion and source terms*, Numerical Methods for Advection-Diffusion Problems 45 (1993) 117–138.

[109] D. W. Peaceman, J. H. H. Rachford, *The numerical solution of parabolic and elliptic differential equations*, Journal of the Society for Industrial and Applied Mathematics 3 (1) (1955) 28–41. `doi:10.1137/0103003`.

[110] P. Chidyagwai, B. Rivière, *Numerical modelling of coupled surface and subsurface flow systems*, Advances in Water Resources 33 (1) (2010) 92–105. `doi:10.1016/j.advwatres.2009.10.012`.

[111] S. Turek, J. Hron, M. Mádlík, M. Razzaq, H. Wobker, J. F. Acker, *Numerical simulation and benchmarking of a monolithic multigrid solver for fluid-structure interaction problems with application to hemodynamics*, Springer Berlin Heidelberg, 2010, pp. 193–220. `doi:10.1007/978-3-642-14206-2_8`.

[112] N. Hanspal, A. Waghode, V. Nassehi, R. Wakeman, *Development of a predictive mathematical model for coupled Stokes/Darcy flows in cross-flow membrane filtration*, Chemical Engineering Journal 149 (1–3) (2009) 132 – 142. `doi:10.1016/j.cej.2008.10.012`.

[113] J. Šukys, S. , Mishra, Ch. Schwab, *Static load balancing for multilevel Monte Carlo finite volume solvers*, in: Parallel Processing and Applied Mathematics, Springer Berlin Heidelberg, 2012, pp. 245–254. `doi:10.1007/978-3-642-31464-3_25`.

[114] D. Drzisga, B. Gmeiner, U. Rüde, R. Scheichl, B. Wohlmuth, *Scheduling massively parallel multigrid for multilevel Monte Carlo methods*, SIAM Journal on Scientific Computing 39 (5) (2017) S873–S897. `doi:10.1137/16M1083591`.

[115] M. A. Celia, E. T. Bouloutas, R. L. Zarba, *A general mass-conservative numerical solution for the unsaturated flow equation*, Water Resources Research 26 (7) (1990) 1483–1496. `doi:10.1029/WR026i007p01483`.

[116] P. Forsyth, Y. Wu, K. Pruess, *Robust numerical methods for saturated-unsaturated flow with dry initial conditions in heterogeneous media*, Advances in Water Resources 18 (1) (1995) 25 – 38. `doi:10.1016/0309-1708(95)00020-J`.

[117] T. Arbogast, M. F. Wheeler, *A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media*, SIAM Journal on Numerical Analysis 33 (4) (1996) 1669–1687. `doi:10.1137/S0036142994266728`.

[118] R. Eymard, M. Gutnic, D. Hilhorst, *The finite volume method for Richards equation*, Computational Geosciences 3 (3) (1999) 259–294. `doi:10.1023/A:1011547513583`.

[119] I. S. Pop, *Error estimates for a time discretization method for the Richards' equation*, Computational Geosciences 6 (2) (2002) 141–160. `doi:10.1023/A:1019936917350`.

[120] C. S. Woodward, C. N. Dawson, *Analysis of expanded mixed finite element methods for a nonlinear parabolic equation modeling flow into variably saturated porous media*, SIAM Journal on Numerical Analysis 37 (3) (2000) 701–724. `doi:10.1137/S0036142996311040`.

[121] A. Mantoglou, L. W. Gelhar, *Stochastic modeling of large-scale transient unsaturated flow systems*, Water Resources Research 23 (1) (1987) 37–46. `doi:10.1029/WR023i001p00037`.

[122] D. Russo, I. Russo, A. Laufer, *On the spatial variability of parameters of the unsaturated hydraulic conductivity*, Water Resources Research 33 (5) (1997) 947–956. `doi:10.1029/96WR03947`.

[123] D. Zhang, *Nonstationary stochastic analysis of transient unsaturated flow in randomly heterogeneous media*, Water Resources Research 35 (4) (1999) 1127–1141. `doi:10.1029/1998WR900126`.

[124] S. C. Iden, W. Durner, *Free-form estimation of the unsaturated soil hydraulic properties by inverse modeling using global optimization*, Water Resources Research 43 (7). `doi:10.1029/2006WR005845`.

[125] M. T. van Genuchten, *A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils*, Soil Science Society of America Journal 44 (1980) 892–898. `doi:10.2136/sssaj1980.03615995004400050002x`.

[126] Y. Mualem, *A new model for predicting the hydraulic conductivity of unsaturated porous media*, Water Resources Research 12 (3) (1976) 513–522. `doi:10.1029/WR012i003p00513`.

[127] D. Zhang, Z. Lu, *Stochastic analysis of flow in a heterogeneous unsaturated-saturated system*, Water Resources Research 38 (2) (2002) 10–15. `doi:10.1029/2001WR000515`.

[128] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Mathematics of Computation 31 (138) (1977) 333–390. `doi:10.2307/2006422`.

[129] A. Brandt, D. Ron, *Multigrid solvers and multilevel optimization strategies*, in: Multilevel optimization in VLSICAD, Springer, 2003, pp. 1–69. `doi:10.1007/978-1-4757-3748-6_1`.

[130] R. E. Bank, H. D. Mittelmann, *Continuation and multi-grid for nonlinear elliptic systems*, in: W. Hackbusch, U. Trottenberg (Eds.), Multigrid Methods II, Springer Berlin Heidelberg, 1986, pp. 23–37.

[131] T. F. C. Chan, H. B. Keller, *Arc-Length continuation and multigrid techniques for nonlinear elliptic eigenvalue problems*, SIAM Journal on Scientific and Statistical Computing 3 (2) (1982) 173–194. `doi:10.1137/0903012`.

[132] N. Dinar, H. B. Keller, *Computations of Taylor vortex flows using multigrid continuation methods*, in: C. C. Chao, S. A. Orszag, W. Shyy (Eds.), Recent Advances in Computational Fluid Dynamics, Springer Berlin Heidelberg, Berlin, Heidelberg, 1989, pp. 191–262.

[133] W. Rheinboldt, *Solution fields of nonlinear equations and continuation methods*, SIAM Journal on Numerical Analysis 17 (2) (1980) 221–237. `doi:10.1137/0717020`.

[134] R. H. Brooks, A. T. Corey, *Hydraulic properties of porous media*, Hydrology Papers, Colorado State University 24 (1964) 37.

[135] C. T. Miller, G. A. Williams, C. T. Kelley, M. D. Tocci, *Robust solution of Richards' equation for nonuniform porous media*, Water Resources Research 34 (10) (1998) 2599–2610. `doi:10.1029/98WR01673`.

[136] O. Ippisch, H. Vogel, P. Bastian, *Validity limits for the van Genuchten-Mualem model and implications for parameter estimation and numerical simulation*, Advances in Water Resources 29 (12) (2006) 1780–1789. `doi:10.1016/j.advwatres.2005.12.011`.

[137] S. Sakamoto, R. G. Ghanem, *Simulation of multi-dimensional non-gaussian non-stationary random fields*, Probabilistic Engineering Mechanics 17 (2) (2002) 167 – 176. `doi:10.1016/S0266-8920(01)00037-6`.

[138] K. K. Phoon, H. W. Huang, S. T. Quek, *Simulation of strongly non-Gaussian processes using Karhunen-Loeve expansion*, Probabilistic Engineering Mechanics 20 (2) (2005) 188–198. `doi:10.1016/j.probengmech.2005.05.007`.

[139] F. Muller, P. Jenny, D. W. Meyer, *A spatially and temporally adaptive solution of Richards' equation*, Advances in Water Resources 29 (4) (2006) 525 – 545. `doi:10.1016/j.advwatres.2005.06.008`.

[140] P. A. Lott, H. F. Walker, C. S. Woodward, U. M. Yang, *An accelerated Picard method for nonlinear systems related to variably saturated flow*, Advances in Water Resources 38 (2012) 92–101. `doi:10.1016/j.advwatres.2011.12.013`.

[141] S. Mehl, *Use of Picard and Newton iteration for solving nonlinear ground water flow equations*, Groundwater 44 (4) (2006) 583–594. `doi:10.1111/j.1745-6584.2006.00207.x`.

[142] J. E. Jones, C. S. Woodward, *A Newton Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems*, Advances in Water Resources 24 (7) (2001) 763 – 774. `doi:10.1016/S0309-1708(00)00075-0`.

[143] G. Juncu, A. Nicola, C. Popa, *Nonlinear multigrid methods for numerical solution of the variably saturated flow equation in two space dimensions*, Transport in Porous Media 91 (1) (2012) 35–47. `doi:10.1007/s11242-011-9831-9`.

[144] J.-G. Caputo, Y. A. Stepanyants, *Front solutions of Richards' equation*, Transport in Porous Media 74 (1) (2008) 1–20. `doi:10.1007/s11242-007-9180-x`.

[145] Y. Zha, J. Yang, L. Yin, Y. Zhang, W. Zeng, L. Shi, *A modified Picard iteration scheme for overcoming numerical difficulties of simulating infiltration into dry soil*, Journal of Hydrology 551 (2017) 56 – 69. `doi:10.1016/j.jhydrol.2017.05.053`.

[146] C. Bierig, A. Chernov, *Convergence analysis of multilevel Monte Carlo variance estimators and application for random obstacle problems*, Numerische Mathematik 130 (4) (2015) 579–613. `doi:10.1007/s00211-014-0676-3`.

[147] F. Muller, P. Jenny, D. W. Meyer, *Multilevel Monte Carlo for two phase flow and Buckley-Leverett transport in random heterogeneous porous media*, Journal of Computational Physics 250 (2013) 685–702. `doi:10.1016/j.jcp.2013.03.023`.

[148] S. Pope, *Turbulent Flows*, Cambridge University Press, 2000. `doi:10.1017/CBO9780511840531`.

[149] Z. Zhang, W. Zhang, Z. J. Zhai, Q. Y. Chen, *Evaluation of various turbulence models in predicting airflow and turbulence in enclosed environments by CFD: Part 2-Comparison with experimental data from literature*, HVAC&R Research 13 (6) (2007) 871–886. `doi:10.1080/10789669.2007.10391460`.

[150] H. Xiao, P. Cinnella, *Quantification of model uncertainty in RANS simulations: A review*. URL `http://arxiv.org/abs/1806.10434`

[151] K. Duraisamy, G. Iaccarino, H. Xiao, *Turbulence modeling in the age of data* (2018) 1–23. URL `http://arxiv.org/abs/1804.00183`

[152] L. Margheri, M. Meldi, M. V. Salvetti, P. Sagaut, *Epistemic uncertainties in RANS model free coefficients*, Computers and Fluids 102 (2014) 315–335. `doi:10.1016/j.compfluid.2014.06.029`.

[153] W. Edeling, P. Cinnella, R. Dwight, H. Bijl, *Bayesian estimates of parameter variability in the $k - \varepsilon$ turbulence model*, Journal of Computational Physics 258 (2014) 73 – 94. `doi: 10.1016/j.jcp.2013.10.027`.

[154] S. H. Cheung, T. A. Oliver, E. E. Prudencio, S. Prudhomme, R. D. Moser, *Bayesian uncertainty analysis with applications to turbulence modeling*, Reliability Engineering and System Safety 96 (9) (2011) 1137–1149. `doi:10.1016/j.ress.2010.09.013`.

[155] M. Emory, R. Pecnik, G. Iaccarino, *Modeling structural uncertainties in Reynolds-Averaged computations of shock/boundary layer interactions*, AIAA paper 479 (2011) 1 – 16.`doi:10.2514/6.2011-479`.

[156] C. Gorlé, G. Iaccarino, *A framework for epistemic uncertainty quantification of turbulent scalar flux models for Reynolds-averaged Navier-Stokes simulations*, Physics of Fluids 25 (5). `doi:10.1063/1.4807067`.

[157] H. Xiao, J.-L. Wu, J.-X. Wang, R. Sun, C. Roy, *Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier-Stokes simulations: A data-driven, physics-informed Bayesian approach*, Journal of Computational Physics 324 (Supplement C) (2016) 115 – 136. `doi:10.1016/j.jcp.2016.07.038`.

[158] C. Gorlé, J. Larsson, M. Emory, G. Iaccarino, *The deviation from parallel shear flow as an indicator of linear eddy-viscosity model inaccuracy*, Physics of Fluids 26 (5). `doi:10.1063/1.4876577`.

[159] E. Dow, Q. Wang, *Quantification of structural uncertainties in the $k - \omega$ turbulence model*, AIAA Paper (2011) 2011–1762.`doi:10.2514/6.2011-1762`.

[160] M. Pisaroni, F. Nobile, P. Leyland, *A continuation multilevel Monte Carlo (C-MLMC) method for uncertainty quantification in compressible inviscid aerodynamics*, Computer Methods in Applied Mechanics and Engineering 326 (2017) 20 – 50. `doi:10.1016/j.cma.2017.07.030`.

[161] M. Leschziner, *Statistical turbulence modelling for fluid dynamics - demystified: An introductory text for graduate engineering students*, Imperial College Press, 2015. `doi:10.1142/p997`.

[162] B. E. Launder, B. I. Sharma, *Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc*, Letters in Heat and Mass Transfer 1 (2) (1974) 131–137. `doi:10.1016/0094-4548(74)90150-7`.

[163] D. C. Wilcox, *Turbulence Modelling for CFD*, DCW industries, 1993.

[164] E. Dow, Q. Wang, *Uncertainty quantification of structural uncertainties in RANS simulations of complex flows*, 20th AIAA Computational Fluid Dynamics Conference, Fluid Dynamics and Co-located Conferences (2011) 2011 – 3865.`doi:10.2514/6.2011-3865`.

[165] C. Soize, *A nonparametric model of random uncertainties for reduced matrix models in structural dynamics*, Probabilistic Engineering Mechanics 15 (3) (2000) 277 – 294. `doi: 10.1016/S0266-8920(99)00028-4`.

[166] C. Soize, *Random matrix theory for modeling uncertainties in computational mechanics*, Computer Methods in Applied Mechanics and Engineering 194 (12) (2005) 1333 – 1366. `doi:10.1016/j.cma.2004.06.038`.

[167] C. Soize, *Maximum entropy approach for modeling random uncertainties in transient elastodynamics*, The Journal of the Acoustical Society of America 109 (5) (2001) 1979–1996. `doi:10.1121/1.1360716`.

[168] C. Soize, *A comprehensive overview of a non-parametric probabilistic approach of model uncertainties for predictive models in structural dynamics*, Journal of Sound and Vibration 288 (3) (2005) 623 – 652. `doi:10.1016/j.jsv.2005.07.009`.

[169] J. Guilleminot, C. Soize, R. G. Ghanem, *Stochastic representation for anisotropic permeability tensor random fields*, International Journal for Numerical and Analytical Methods in Geomechanics 36 (13) (2012) 1592–1608. `doi:10.1002/nag.1081`.

[170] D. Xiu, G. E. Karniadakis, *The Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing 24 (2) (2002) 619–644. `doi:10.1137/S1064827501387826`.

[171] OpenFOAM, available from http://www.openfoam.org.

[172] S. Patankar, D. Spalding, *A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows*, International Journal of Heat and Mass Transfer 15 (10) (1972) 1787 – 1806. `doi:10.1016/0017-9310(72)90054-3`.

[173] B. Basara, S. Jakirlic, *A new hybrid turbulence modelling strategy for industrial CFD*, International Journal for Numerical Methods in Fluids 42 (1) (2003) 89–116. `doi:10.1002/fld.492`.

[174] A. Huser, S. Biringen, *Direct numerical simulation of turbulent flow in a square duct*, Journal of Fluid Mechanics 257 (1993) 65–95. `doi:10.1063/1.4922087`.

[175] M. Breuer, N. Peller, C. Rapp, M. Manhart, *Flow over periodic hills -Numerical and experimental study in a wide range of Reynolds numbers*, Computers and Fluids 38 (2) (2009) 433 – 457. `doi:10.1016/j.compfluid.2008.05.002`.

[176] S. Banerjee, R. Krahl, F. Durst, Ch. Zenger, *Presentation of anisotropy properties of turbulence, invariants versus eigenvalue approaches*, Journal of Turbulence 8 (2007) N32. `doi:10.1080/14685240701506896`.

[177] R. L. Thompson, L. E. B. Sampaio, F. A. de Bragança Alves, L. Thais, G. Mompean, *A methodology to evaluate statistical errors in DNS data of plane channel flows*, Computers & Fluids 130 (2016) 1 – 7. `doi:10.1016/j.compfluid.2016.01.014`.

[178] T. J. Dodwell, C. Ketelsen, R. Scheichl, A. L. Teckentrup, *A hierarchical multilevel Markov chain Monte Carlo algorithm with applications to uncertainty quantification in subsurface flow*, SIAM/ASA Journal on Uncertainty Quantification 3 (1) (2015) 1075–1108. `doi:10.1137/130915005`.

[179] C. Rodrigo, P. Salinas, F. J. Gaspar, F. J. Lisbona, *Local Fourier analysis for cell-centered multigrid methods on triangular grids*, Journal of Computational and Applied Mathematics 259 (2014) 35 – 47. `doi:10.1016/j.cam.2013.03.040`.

[180] C. Rodrigo, F. J. Gaspar, X. Hu, L. T. Zikatanov, *Stability and monotonicity for some discretizations of the Biot's consolidation model*, Computer Methods in Applied Mechanics and Engineering 298 (2016) 183 – 204. `doi:10.1016/j.cma.2015.09.019`.

[181] P. Luo, C. Rodrigo, F. J. Gaspar, C. W. Oosterlee, *Monolithic multigrid method for the coupled Stokes flow and deformable porous medium system*, Journal of Computational Physics 353 (2018) 148 – 168. `doi:10.1016/j.jcp.2017.09.062`.

[182] G. Geraci, M. S. Eldred, G. Iaccarino, *A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications*, 19th AIAA Non-Deterministic Approaches Conference, (2017). `doi:10.2514/6.2017-1951`.

Curriculum Vitae

Prashant Kumar

19-11-1990   Born in Ranchi, India

Education

2009-2013   Bachelor of Science in Exploration Geophysics
Indian Institute of Technology Kharagpur, India

2013-2014   Master of Science in Exploration Geophysics
Indian Institute of Technology Kharagpur, India

2015-2019   PhD in Applied Mathematics
Delft University of Technology, Delft, The Netherlands
Thesis: Multilevel solvers for stochastic fluid flows
Promotors: Dr. R. P. Dwight, Prof. dr. ir. C. W. Oosterlee

List of Publications

Publications in refereed journals

1. Kumar, P., Rodrigo, C., Gaspar, F. J. & Oosterlee, C. W. (2019). **On local Fourier analysis of multigrid methods for PDEs with jumping and random coefficients.** *SIAM Journal on Scientific Computing*, 41(3), A1385-A1413.

2. Kumar, P., Rodrigo, C., Gaspar, F. J. & Oosterlee, C. W. (2019). **A parametric acceleration of multilevel Monte Carlo convergence for nonlinear variably saturated flows.** *Preprint arXiv:1903.08741.* (Submitted for publication)

3. Kumar, P., Luo, P., Gaspar, F. J. & Oosterlee, C. W. (2018). **A multigrid multilevel Monte Carlo method for transport in the Darcy-Stokes system.** *Journal of Computational Physics*, 371, 382-408.

4. Kumar, P., Schmelzer, M. & Dwight, R. P. (2018). **Stochastic turbulence modeling in RANS simulations via Multilevel Monte Carlo method.** *Preprint arXiv:1811.00872.* (Submitted for publication)

5. Kumar, P., Oosterlee, C. W. & Dwight, R. P. (2017). **A multigrid Multilevel Monte Carlo method using high-order finite-volume scheme for lognormal diffusion problem.** *International Journal for Uncertainty Quantification*, 7(1), 57-81.

List of attended conferences with presentations

Oral presentation

1. **Fifteenth International Conference on Mathematics and its Application**, Jaca, Spain, September 2018.

2. **Workshop on Multilevel and Multigrid Methods** (invited), SIAM student chapter TU Delft, the Netherlands, May 2018.

3. **The Computational Mathematics Aspects of Porous Media and Fluid Flow** (invited), Lorentz centre workshop, Leiden, the Netherlands, May 2018.

4. **SIAM conference on Uncertainty Quantification**, Los Angeles, USA, April 2018.

5. **European Conference on Numerical Mathematics (ENUMATH)**, Voss, Norway, September, 2017.

6. **Computational Sciences for Future Energy Conference**, Utrecht, the Netherlands, September, 2017.

7. **Frontiers of Uncertainty Quantification in Engineering (FrontUQ)**, Munich, Germany, September, 2017.

8. **Computational Sciences for Future Energy Conference**, Eindhoven, the Netherlands, October, 2016.

9. **SIAM conference on Uncertainty Quantification**, Lausanne, Switzerland, April 2016.

10. **Copper Mountain Conference on Iterative Methods**, Colorado, USA, March 2016.

Poster presentation

1. **The 43rd Woudschoten Conference, Zeist** *(second prize out of 35 posters)*, the Netherlands, October 2018.

2. **The 42nd Woudschoten Conference, Zeist**, the Netherlands, October 2017.

3. **Kick-off DCSE Event entitled Future Computing**, Delft, the Netherlands, May 2017.

Acknowledgements

This thesis could not have been completed without the support of many people. Here I would like to acknowledge some of them in particular.

First and foremost, I would like to express my sincere gratitude to my promotors Prof. Kees Oosterlee and Dr. Richard Dwight for their invariable support and encouragement throughout the PhD journey.

Kees, thanks for providing me with such an amazing opportunity to work at CWI. Your supervision over the past four years has deeply influenced me both academically and personally. Thanks for taking out time for weekly meetings and reviewing my progress in detail ensuring I stayed on schedule. Your sharp bits of advice about scientific writing significantly improved my ability to put complex ideas into comprehensible sentences and formulas. Thanks for being prompt with proofreading my articles and for your remarkable help in writing the dissertation. Your strong support and guidance for starting the company are very much appreciated.

Richard, it was a pleasure working with you and I am forever grateful to you for agreeing to be my daily supervisor, even before knowing me. Coming from a geophysics background, I owe my understanding about many topics in aerodynamics to you. Thank you for allowing me to work in a way that was most fitting to me while I was working on two very different projects around the same time. Thanks for deciding to visit CWI on Thursdays, saving me numerous trips to Delft. I really enjoyed our interesting after-squash discussions at Oerknal.

I would also like to thank Prof. Francisco Gaspar from the University of Zaragoza, Spain. We collaborated intensively during your two-year stay at CWI and also during my month-long visit at Zaragoza. From you, I learned a lot about discretization, multigrid and red wines. I have a lot of happy memories from Zaragoza exploring tapas bars, wine tasting and teaching me how to cook a Spanish omelette. A very special thanks to Carmen for being a wonderful host during my stay at Zaragoza. Thanks for taking out time from your busy teaching schedule to review my articles.

Laurent, it was a pleasure sharing the office with you for the last four years. I appreciate your help on numerous occasions and big thanks for helping me translate my thesis summary. You went all the way to reading a turbulence-modeling textbook in Dutch to improve the summary translation. I always enjoyed our ping-pong sessions and I am very impressed by the way you managed to calibrate your game with your non-conventional racket grip!

Daan, thanks for your efforts in the role of the group leader and providing a supportive environment for researchers in the group. Benjamin and Svetlana, thanks for organizing CRUQS and SC seminar from which I learned a lot and got a few good ideas for my research as well. Very special gratitude goes out to Nada at CWI and Colette at TU Delft for their enormous help with administrative tasks. I am also grateful to the following CWI staffs: Bikkie, Duda, and Minnie for helping me out on many occasions.

Yous and Anastasia, last few months have been the most exciting time working with you guys starting-up our company. I look forward to many more adventures together in the future. Yous, thanks for lending me your awesome guitar, it was a big step-up from Ukulele. Special thanks go out to Krzysztof, Ki Wai and Sangeetika for the many fun evenings, nights, and weekends, which gave me much needed distractions from my PhD research. Alvaro, we had good times in Amsterdam, thank you for being a brilliant squash rival and a teacher. *Zatte* does not taste the same without you!