Technische Universität München

Fakultät für Mathematik
Professur für Wissenschaftliches Rechnen (Prof. Dr. Elisabeth Ullmann)

Variance reduction with multilevel estimators

Daniel Schaden

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Michael Ulbrich

Prüfer der Dissertation:

1. Prof. Dr. Elisabeth Ullmann
2. Prof. Dr. Benjamin Peherstorfer
3. Prof. Dr. Stefan Vandewalle

Die Dissertation wurde am 20.01.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 05.04.2021 angenommen.

Titel in deutscher Sprache:
Varianzreduzierung mit Multilevel–Schätzern

Zusammenfassung:
Diese Dissertation besteht aus zwei Teilen, die sich beide mit partiellen Differentialgleichungen mit zufälligen Koeffizienten befassen, welche bei der Uncertainty Quantification auftreten. Ziel der Arbeit ist es kosteneffiziente Schätzer zu konstruieren indem Diskretisierungen der partiellen Differentialgleichungen mit unterschiedlichen Genauigkeiten kombiniert werden.
Im ersten Teil stellen wir eine multilevel Varianzreduktionstechnik vor, um den Erwartungswert einer relevanten Größe zu schätzen. Außerdem analysieren wir diese. Die Hauptidee besteht darin, die Schätzung als verallgemeinertes lineares Kleinste-Quadrate-Problem neu zu formulieren und den zugehörigen multilevel besten linearen erwartungstreuen Schätzer herzuleiten. Wichtig ist, dass dieser Schätzer bei einer Hierarchie von Modellen anwendbar ist. In einem weiteren Schritt betrachten wir die Berechnungskosten der Samples und konstruieren einen sample allocation optimal best linear unbiased estimator (SAOB). Dieser Schätzer erreicht die kleinste Varianz in der Klasse der linearen erwartungstreuen Schätzer mit einem vorgeschriebenen Rechenbudget. Somit verbessert der SAOB bestehende Methoden wie Monte Carlo, Multilevel Monte Carlo und Multifidelity Monte Carlo. Man kann zeigen, dass die Komplexität des SAOB asymptotisch optimal ist für lineare Kombinationen von Samples aus Modelldiskretisierungen, die gegen die exakte relevante Modellgröße konvergieren. Es ist jedoch schwierig, explizite Ausdrücke für die Komplexität des implizit definierten SAOB zu erhalten. Aus diesem Grund führen wir die neuen Richardson–Extrapolations–Schätzer ein und analysieren sie um die Kosten des SAOB nach oben abzuschätzen. Interessanterweise ist der Richardson–Extrapolations–Schätzer eine Verallgemeinerung des Multilevel–Monte–Carlo–Schätzers.
Im zweiten Teil entwickeln wir einen Multilevel–Monte–Carlo–Schätzer für ein risikoneutrales Optimalsteuerungsproblem mit deterministischer Kontrolle. Die Grundidee besteht darin, die Multilevel Monte Carlo Diskretisierung vom Erwartungswert in der Zielfunktion auf die deterministische Kontrolle zu verschieben. Dies liefert eine Folge konvexer Optimierungsprobleme. Wir zeigen, dass dies ähnlich wie bei der normalen Multilevel Monte Carlo Methode die Varianz des Schätzers der optimale Kontrolle verringert. Im Gegensatz zu alternativen Methoden in der Literatur, beispielsweise stochastischen Optimierungsmethoden, ist keine Auswahl der Schrittweite erforderlich. Darüber hinaus kann die Konvergenzanalyse des neuen Ansatzes mit klassischen Werkzeugen aus der numerischen Analyse durchgeführt werden. Wir verifizieren die Hauptergebnisse dieser Arbeit numerisch unter Verwendung einer elliptischen partiellen Differentialgleichung mit zufälligen Koeffizienten.

Abstract

This thesis has two parts, both concerned with partial differential equations with random coefficients arising in uncertainty quantification. The goal of the thesis is to construct cost-efficient estimators by combining discretizations of the partial differential equations with different accuracies.
In the first part we introduce and analyse a multilevel variance reduction technique to estimate the expectation of a quantity of interest. The main idea is to reformulate the estimation as a generalized linear least squares problem and derive the associated multilevel best linear unbiased estimator. Importantly, this estimator can work with a hierarchy of models. In a further step we consider the computational cost for a sample and construct a sample allocation optimal best linear unbiased estimator (SAOB). This estimator achieves the smallest variance in the class of linear unbiased estimators given a prescribed computational budget. Thus, the SAOB improves upon existing methods like Monte Carlo, Multilevel Monte Carlo and Multifidelity Monte Carlo. We show that the complexity of the SAOB is asymptotically optimal for linear combinations of samples from model discretizations which converge to the exact model output quantity of interest. However, explicit expressions for the complexity of the implicitly defined SAOB are difficult to obtain. For this reason we introduce and analyse the novel Richardson extrapolation estimators that allow us to upper bound the cost of the SAOB. Interestingly, the Richardson extrapolation estimator is a generalisation of the Multilevel Monte Carlo estimator.
In the second part, we develop a Multilevel Monte Carlo estimator for a risk neutral optimal control problem with deterministic control. The basic idea is to push the Multilevel Monte Carlo discretization from the mean in the cost functional to the deterministic control which leads to a sequence of convex optimization problems. We show that, similar to standard Multilevel Monte Carlo, the variance of the estimator for the optimal control is reduced. In contrast to alternative methods in the literature, for example, stochastic optimization methods, no step size selection is required. In addition, the convergence analysis of the new approach can be carried out with classical tools from numerical analysis. We numerically verify the main results of this thesis using an elliptic partial differential equation with random coefficients.

Acknowledgements

I express my gratitude for my advisor Elisabeth Ullmann for her knowledge, experience, patience and support throughout my doctoral research. Her suggestions and remarks tremendously improved my writing and presentation skills and thus improved the papers I contributed to as well as my presentations. Without her feedback the quality and scope of my research and this thesis would not be the same.
During my research I worked with the graduate students of the IGDK1754, in particular with Sören Behr, Sebastian Engel, Dominik Hafemeyer, Gernot Holler (Graz), Sandra Marschke (Graz), Johannes Milz, Christian Münch and Daniel Walter (Linz). I would like to thank all of them for their time, help and suggestions.
I would like to also thank Jonas Latz for his helpful remarks regarding Bayesian inversion for the Helmholtz paper. I am grateful for the finite element code provided by Michael Ulbrich. A modification of this code was used for some numerical experiments in this thesis.
I thank Barbara Wohlmuth and Daniel Drzisga for the use of the local cluster and in particular, I would like to thank Laura Scarabosio (Nijmegen) for her help with numerical experiments. I would also like to thank my co-workers Mario Teixeira Parente and Fabian Wagner for helpful discussions.
During my school education I had the pleasure to take part in the Robotics-AG at the Gymnasium Weingarten and I express my gratitude to Hansjörg Stengel. I would also like to thank Dominik Meidner and Boris Vexler for their help in regards to my Bachelor's thesis. I also thank Manfred Liebmann (Graz) for his help during my Master's thesis and his insights into high performance computing and machine learning. I would like to thank Karl Kunisch (Graz) for his comments on my work and the pleasant research stay in Graz.
I acknowledge the help of Jenny Radeck, Diane Clayton-Winter and Vanessa Peinhart (Graz) for organizational aspects of my research.
Finally, I thank my parents Iris and Thomas and my brothers Tobias and Benjamin for their help and support throughout my studies and research in Munich.

I assure the single handed composition of this doctoral thesis only supported by declared resources.

Garching bei München, January 10th, 2021

.....
Daniel Schaden

Publications by the author

Parts of this thesis contain excepts from articles that are published or submitted and under review. These articles are part of the doctoral research and this thesis. Daniel Schaden is the main author of:

[126] D. Schaden and E. Ullmann. On Multilevel Best Linear Unbiased Estimators. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):601–635, 2020

[125] D. Schaden and E. Ullmann. Asymptotic analysis of multilevel best linear unbiased estimators, *arXiv*:2012.03658, 2020. *submitted*

Chapter 4 contains results of [126]. Chapter 5 and Chapter 6 contain and extend results from both [125] and [126]. Excerpts of [125, 126] may also be contained in other chapters.

Daniel Schaden also researched Bayesian inversion. The following article is not contained in this thesis:

[49] S. Engel, D. Hafemeyer, C. Münch, and D. Schaden. An application of sparse measure valued Bayesian inversion to acoustic sound source identification. *Inverse Problems*, 35(7):075005, 2019

Garching bei München, January 10th, 2021

.....
Daniel Schaden

Contents

1 Introduction **13**
1.1 Organization of the thesis 15

2 Partial differential equations with random coefficients **19**
2.1 Probability Theory 19
2.2 Karhunen–Loève expansion 28
2.3 Elliptic partial differential equation and discretization 31

3 Estimation and variance reduction **41**
3.1 Sampling based estimation 43
3.2 Monte Carlo 45
3.3 Control Variates 48
3.4 Multifidelity Monte Carlo 53
3.5 Approximate Control Variates 63
3.6 Multilevel Monte Carlo 67
3.7 Other multilevel methods 71

4 Multilevel best linear unbiased estimators **73**
4.1 Estimation as linear regression 74
4.2 Lower variance bound 85
4.3 Linear subspace formulation 88
4.4 Comparison of linear unbiased estimators 90
4.5 Numerical experiments 101

5 Sample allocation optimal BLUE **109**
5.1 Ideal sample allocation optimal BLUE 109
5.2 First BLUE, then sample allocation 113
5.3 First sample allocation, then BLUE 122
5.4 Characterisation of the set of minimizers 129

6 Asymptotics of the SAOB **139**
6.1 Asymptotic analysis for linear unbiased estimators 140
6.2 Richardson Extrapolation Estimator 146
6.3 Lower bounds on the complexity 157
6.4 Numerical experiments with explicit expansions 162
6.5 Numerical experiments with an elliptic PDE 167

7 A multilevel approach for the risk neutral optimal control problem **175**
7.1 The risk neutral optimal control problem 176
7.2 Monte Carlo discretization 182
7.3 Multilevel Monte Carlo for the control 189
7.4 Numerical experiments 201

8 Conclusion and outlook **207**

Bibliography **215**

List of Figures **225**

List of Tables **226**

List of Symbols **227**

List of Abbreviations **232**

Chapter 1
Introduction

Uncertainty quantification is an important branch of Mathematics. Uncertainties arise from measurement errors, unobservable or only partially known data as well as model errors which are incorporated into physical and mathematical models [54, 130]. This is done by assuming that the input of the model is random and takes on a possibly infinite set of values with a prescribed probability. The input can be obtained from real life measurements, generated samples from a simple distribution or the result of another model. The response or output of the model then assumes a possibly infinite set of values and follows a probability distribution, which allows us to study the model under different configurations. This field of study is often called forward uncertainty quantification to distinguish it from the more challenging inverse uncertainty quantification, where the output is given and the input (distribution) has to be reconstructed, see e.g. [37, 129] or [130, Chapter 6].
The infinite set of model outputs are difficult to examine and thus often collapsed in a meaningful way into fewer values. The most important statistic is the expected value or average output of the model. Other important statistics include the variance, moments, quantiles or risk measures, see [110] or [127, Section 6] for the latter. The models and the probability distribution of the inputs are often complicated such that there is no analytic expression for the probability distribution of the output. Hence, the statistics of the output have to be estimated or approximated with a numerical scheme. A well–known technique for the estimation is the Monte Carlo method [77, 79, 118] which samples from the input distribution or uses representative samples from real life measurements. The model is then simulated and the computed or measured outputs are averaged to obtain an approximation for the expectation.
There are numerous applications of the above approach. Classical use cases for probability theory and estimation include financial products, where we want to compute the expected return, estimate the risk of a default or optimize a portfolio with respect to (w.r.t.) some metric, see [59]. Similar methods are used to analyse related fields like stochastic games and gambling. Machine learning [69, 99, 132] is another use case and estimation is specifically used to train regression models like neural networks or kernel methods. The models are trained for complicated tasks, for example image classification, face recognition, targeted advertising, knowledge discovery or reinforcement learning for board games. Other examples include groundwater flow where the composition of the underlying rock layers is not fully known [40, 141]. In this thesis, we concentrate on a randomized version of Poisson's equation which models the stationary temperature profile in a material with unknown heat conductivity coefficient. This is the standard model in forward uncertainty quantification [25, 31, 63, 87, 134]. We assume that the conductivity coefficient is a lognormal random field and we use the Karhunen–Loève expansion [1, Section 3] to sample from it. We now informally describe the methods, motivation and goals of this thesis. Afterwards, we give a brief summary of the contents and main results of each chapter.

Estimation. The Monte Carlo method is an extremely general method that can be used to estimate the mean. The idea is to average the model outputs for multiple inputs

following the same distribution. It relies only on a few weak assumptions, is often easy to implement and does not require any knowledge of the underlying distribution, except for existence of the first moment. Furthermore, the Monte Carlo method does not suffer from "the curse of dimensionality", which is often in contrast to deterministic quadrature rules, see [24][Section 2], [35, Section 5.4] or [46, Section 1]. The curse of dimensionality is a phrase to emphasize that the cost of a method increases rapidly, sometimes even exponential, with its dimension. The generality of the Monte Carlo method and its easy use has the significant downside of not being very cost effective. Indeed, this method often requires a large amount of samples and thus we have to compute the model response very often. Considerable research and methods have been proposed to improve and speed up basic Monte Carlo. An often used term in this context is variance reduction since the cost of the Monte Carlo method is often proportional to the variance of the random model output [59, 118]. All parts in this thesis are geared towards achieving and obtaining a variance reduction with sampling based methods. We mainly focus on the control variate approach and neglect other approaches that modify the sampling process like importance sampling, Markov chain Monte Carlo or Sequential Monte Carlo [28, 48, 54, 118].

Model discretization and variance reduction. Models like Poisson's equation for diffusion processes often require numerical approximations, since the respective solution cannot be computed analytically. This requires us to discretize an infinite dimensional function space and we use the well–known finite element method [21, 29] which approximates this space with a finite number of basis functions. The approximation quality increases if we increase the number of basis functions, hence the costs to obtain an approximate solution also increases. This means that there is an inherent trade off between the accuracy of the solution and the computational costs. Multigrid methods [67, 137] use coarse grids to reduce the effort to solve a linear system on the fine grid. The idea to use coarse models for estimation was used by Heinrich [70] and the Multilevel Monte Carlo approach was analysed by Giles [56, 57]. A control variate approach, where coarse grid levels are used, is the Multifidelity Monte Carlo estimator [106, 107] or the Approximate Control Variate approach [62]. A survey of Multifidelity methods for estimation can be found in [108]. As it turns out, if the model discretization satisfies some cost and variance properties, then the Multilevel Monte Carlo estimator of Giles [56] achieves a substantially smaller asymptotic cost than Monte Carlo. This means the actual model, whose mean we want to approximate, can be estimated much cheaper. This was also verified analytically for the Multifidelity Monte Carlo estimator in [106].

Best linear unbiased estimators and sample allocation. We show that it is helpful to view the estimation of the mean as regression problem. The best linear unbiased estimator, which is a well–known method in Statistics [8, 64, 69, 96, 114, 116, 142] uses a linear combination of samples to estimate a parameter. We systematically develop estimators that combine samples from inaccurate but cheap models with accurate but expensive models. In contrast to Multilevel Monte Carlo methods, which exploits a similar idea to drastically reduce the costs at least in the case of hierarchical models, we emphasize the viewpoint as regression problem. We furthermore optimize the sample allocation that determines the used models and how often we evaluate them. This then leads to an estimator that is cost minimal in the class of linear unbiased estimators. Sample allocation problems are crucial for a good estimator and this was already discussed in [56] and [107] for the respective estimators that allow for a unique sample allocation under mild assumptions.

Optimal control problems. The goal in optimal control problems is to find a control that steers the response of a system towards a prescribed desired state [73, 136]. For example, the temperature inside a material should be close to the desired temperature and we can cool or heat the material only at the boundary. Mathematically speaking, this can be formulated as constrained minimization problem, where the solution is the optimal control. The distinguishing feature is that the response of the system cannot be controlled directly but only indirectly. The conductivity coefficient is often unknown and thus assumed to follow some probability distribution. This problem is a risk neutral optimal control problem and has gained interest in the literature, where different variants and solution methods are discussed [4, 16, 52, 82, 138]. We search for a deterministic control, however, the response of the system is random. Therefore the control is chosen to be close on average to the desired state. We propose a novel variance reduction technique based on the Multilevel Monte Carlo method to solve this minimization problem.

1.1 Organization of the thesis

This thesis is organized in eight chapters, where the first one is the introduction. We list the other chapters with their respective content and objective. Figure 1.1 shows the ordering of the chapters.

Chapter 2: Partial differential equations with random coefficients. In this chapter we introduce concepts needed for forward uncertainty quantification. This includes basic probability theory, where we introduce random variables, their expectations and variances. We further discuss the Karhunen–Loève expansion which is a method to generate samples with values in an infinite dimensional Hilbert space. We conduct numerical experiments with the help of Poisson's equation, which we discuss in the last section of this chapter. We further provide known results for the accuracy of the finite element approximation of the solution.

Chapter 3: Estimation and variance reduction. We present methods to estimate the expectation of a quantity of interest. These methods are used in practice and well known in the literature. We start with the Monte Carlo method and introduce the control variate approach to obtain a variance reduction. Practically implementable control variate approaches are the Multifidelity Monte Carlo and Approximate Control Variates estimator, which improve over standard Monte Carlo in certain circumstances. We further introduce the Multilevel Monte Carlo method, which is another method to reduce the variance. We provide asymptotic results of these estimators for a model sequence converging to the true model and introduce the notion of a lower variance bound. We use these methods as comparison to the best linear unbiased estimator or the SAOB in the following chapters.

Chapter 4: Multilevel best linear unbiased estimators. We present the basic idea behind multilevel best linear unbiased estimators (BLUE) in this chapter. We examine the class of linear estimators that use linear combinations of the samples and are unbiased w.r.t. some linear combination of the mean values. A well–known result is that there exists a best linear unbiased estimator, where best means that the variance is smallest. Importantly, we reformulate the estimation of the mean as linear regression problem which allows us to use the available mathematical literature for least squares problems. The multilevel BLUE is then the (unique) solution of this regression problem, where

the expression "multilevel" refers to different discretization levels of the same quantity of interest. We then examine the estimators of Chapter 3 and give details under what circumstances these are BLUEs.

Chapter 5: Sample allocation optimal BLUE. The BLUE is defined as solution of a regression problem. This regression problem itself depends on the sample allocation, which we study in this chapter. We introduce a budget constraint and the costs for the evaluation of a model group. A model group is a collection of models that we evaluate using the same input sample. The goal is then to select the regression problem and estimator such that the variance is minimized given a fixed budget. We call the resulting estimator the SAOB, which is optimal in the class of linear unbiased estimators. Computing the true SAOB is in general an intractable problem due to integer constraints and thus focus mostly on a relaxed version. We show that the relaxed sample allocation problem has a solution, verify that this solution is in general not unique and a solution can be found that uses at most L model groups. We then proceed and show that first optimizing the sample allocation and then the coefficients is beneficial to remove some assumptions of the previous theorems. The resulting optimization problem is similar to an ℓ_1 minimization problem, which allows us to show that the set of minimizers has a specific structure.

Chapter 6: Asymptotics of the SAOB. This chapter extends the asymptotic analysis of the Multilevel Monte Carlo estimator to the SAOB. The latter estimator is only given implicitly as a minimizer of a convex optimization problem and thus rather difficult to analyse. However, since the SAOB is the linear unbiased estimator with the smallest variance given some prescribed budget, we are able to bound its complexity with explicit complexity bounds from other linear unbiased estimators. In particular, we are able to use the Monte Carlo, the Multifidelity and the Multilevel Monte Carlo estimator. We then introduce Richardson extrapolation for both the mean and the variance to obtain an estimator that, under specific assumptions, has an improved complexity compared to other estimators. The obtained complexity bounds, while not necessarily sharp, are also valid for the SAOB.

Chapter 7: A multilevel approach for the risk neutral optimal control problem. We introduce a multilevel approach for solving optimization problems, which we exemplary apply for a linear quadratic optimization problem. This problem is a risk neutral optimal control problem where we compute a deterministic control such that the systems response is on average close to some prescribed state. A straightforward and naive application of the Multilevel Monte Carlo estimator leads to an ill–posed optimization problem. Our approach is to push the Multilevel Monte Carlo discretization to the deterministic control, hence the name Multilevel Monte Carlo for the control. We verify that this leads to a sequence of well–posed convex optimization problems. Furthermore, this substantially improves the cost that are up to logarithmic factors equal to the cost of the standard Multilevel Monte Carlo estimator.

Chapter 8: Conclusion and outlook. We finish with a conclusion of this thesis and discuss open problems together with possible future research directions.

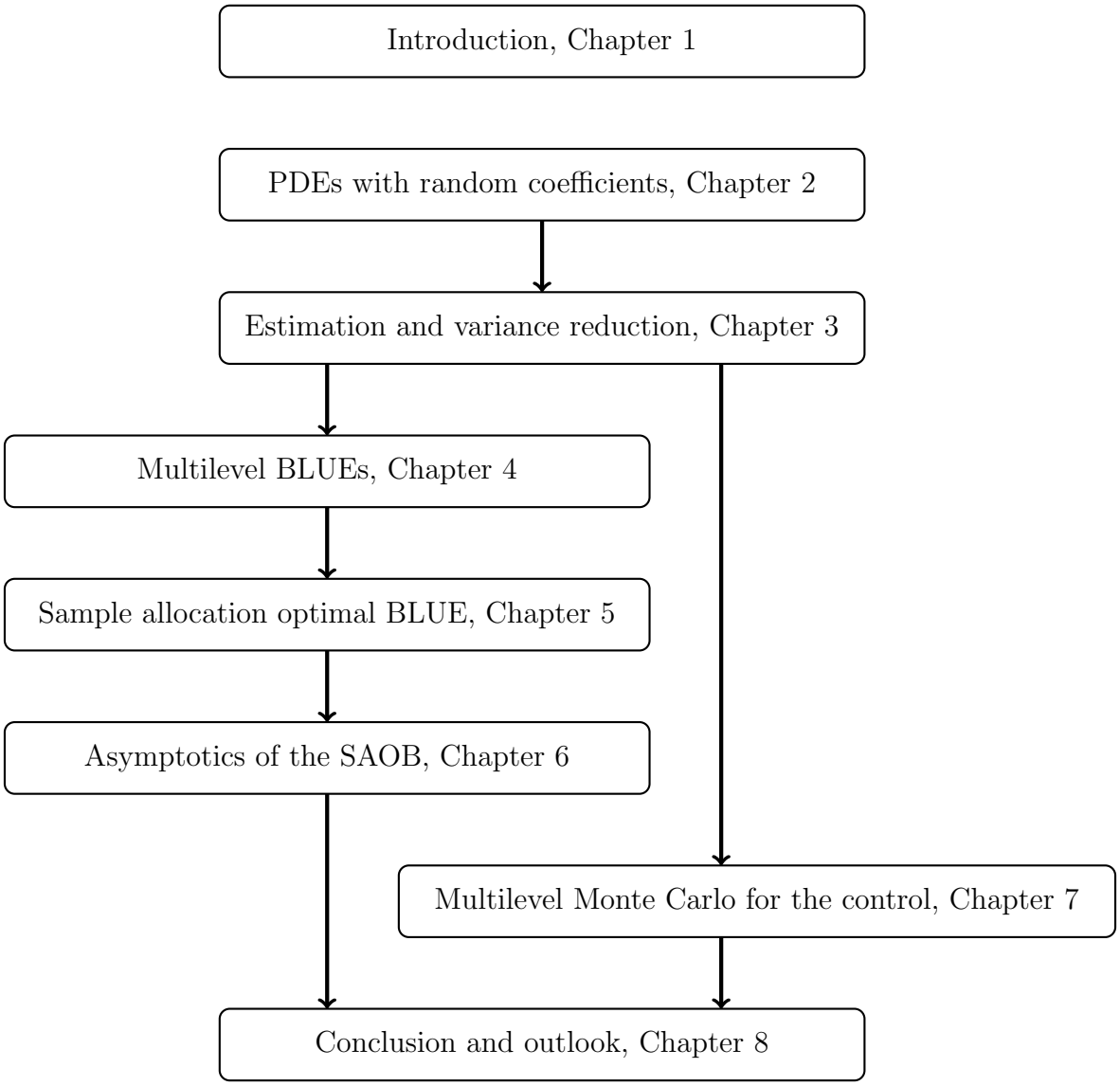

Figure 1.1: Chapters and ordering of this thesis.

Chapter 2

Partial differential equations with random coefficients

In this chapter we describe the basic notation, definition and results needed as foundation of this thesis. Basic in this context means that we describe well–known ideas and methods in mathematics. We lay a common ground to present our results in later chapters. Each section of this chapter contains a short outline of its topic and concepts adapted for this thesis, where proofs of statements are mostly omitted or very short to highlight the main idea. We provide three distinct sections regarding related parts of Forward Uncertainty Quantification.

- **Probability Theory**: Section 2.1 contains a short introduction and repetition of basic notation in probability theory. We explain concepts like random variables, independence, expectation, variance and provide some useful inequalities. Here we use standard definitions from [79]. Further introductions to probability theory can also be found in [6, 10, 77].

- **Karhunen–Loève expansion**: We are interested in random variables that have realizations in a function space and we use the Karhunen–Loève expansion to generate samples, which we describe in Section 2.2. We mainly focus on mean zero Gaussian random fields and provide results for the Whittle–Matérn covariance function. We further discuss some practical methods how to compute and sample from a Karhunen–Loève expansion.

- **Mathematical models**: We provide the mathematical models we are using in thesis in Section 2.3. This is mostly Poisson's equation, which can be used to model the temperature of a material given some heat source. We randomize the conductivity and provide results for the existence and uniqueness of the random solution of the weak Poisson's equation. The finite element method is used to obtain a discretized and thus computable solution which converges to the exact solution with a certain rate.

2.1 Probability Theory

Random variables. The foundation of modern probability theory is a probability space. We provide the standard definition and names for related concepts as well.

Definition 2.1 (Probability space [79, Definition 1.38])**.** The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is a *probability space* if

- the *set of elementary events* Ω is non–empty, $\Omega \neq \emptyset$,
- the *σ-algebra* $\mathcal{F}$ is a suitable subset of the *power set*, $\mathcal{F} \subseteq 2^{\Omega}$,
- the *probability measure* $\mathbb{P}$ is a suitable measure, $\mathbb{P} : \mathcal{F} \to [0, 1]$. ◇

We use the usual definition of σ-algebra, the power set and probability measure. These definitions are available in Klenke [79, Section 1] or Kallenberg [77, Section 1, Section 2]. We denote *(elementary) events* with $\omega \in \Omega$ and also refer to the measure $\mathbb{P}$ as *(probability) distribution*. We call the pair $(\Omega, \mathcal{F})$ *measurable space*. The motivation for Definition 2.1 is to assign every observable event $F \in \mathcal{F}$ a probability between $[0, 1]$ which models the chance of it occurring.
We are typically not interested in the probability of a random event $F \in \mathcal{F}$ but rather on a respective outcome or effect. This is modelled with the help of random variables, which are measurable functions that map events ω to quantities that we are interested in.

Definition 2.2 (Random variable [77, Section 2])**.** Let $(H, \mathcal{G})$ be a measurable space and $Z : \Omega \to H$ a function. Z is *measurable* if the preimage of a measurable set is measurable. Stated formally, we require that for all $G \in \mathcal{G}$

$$Z^{-1}(G) := \{\omega \in \Omega \,|\, Z(\omega) \in G\} \in \mathcal{F}.$$

If Z is measurable on a probability space then Z is called a *random variable*. Then, in case H is a space of vectors, we call Z a *random vector* and if H is a space of functions, we call Z a *random field*. For $H := \mathbb{R}$ we call Z *real–valued*. ◇

The random variable Z allows us to define the probability of certain outcomes in the image space H. For all $G \in \mathcal{G}$ the probability that Z assumes values in G is

$$\mathbb{P}(Z \in G) := \mathbb{P}(\{\omega \in \Omega \,|\, Z(\omega) \in G\}) = \mathbb{P}(Z^{-1}(G)).$$

This expression is well defined since Z is measurable by its definition as random variable. As a consequence we conclude that the triple $(H, \mathcal{G}, \mathbb{P}(Z \in \cdot))$ is a well–defined probability space. The map $\mathbb{P}(Z \in \cdot)$ is called the *pushforward* of $\mathbb{P}$ under Z. We denote this by $Z \sim \mathbb{P}$ and call $\mathbb{P}$ the *distribution* of Z.
The composition of measurable functions is again measurable and thus we are able to compose new random variables in a simpler way. We precisely state this result.

Lemma 2.3 (Composition of measurable functions [79, Theorem 1.80])**.**
Let $(\Omega, \mathcal{F})$, $(H_1, \mathcal{G}_1)$ and $(H_2, \mathcal{G}_2)$ be measurable spaces. Furthermore, let $Z_1 : \Omega \to H_1$ and $Z_2 : H_1 \to H_2$ be measurable functions w.r.t. the respective σ-algebras. Then the composition $Z_2 \circ Z_1 : \Omega \to H_2$ is measurable.

Proof. For $G \in \mathcal{G}_2$ the preimage $Z_2^{-1}(G)$ is measurable w.r.t. $\mathcal{G}_1$ and thus $Z_1^{-1}(Z_2^{-1}(G))$ is measurable w.r.t. $\mathcal{F}$. This concludes the proof. □

An important class of measurable functions are continuous functions. This often allows us to circumvent the rather tedious direct verification of measurability from its definition. We require the notion of a *topological space*, which is formally defined by Klenke [79, Definition 1.20]. A topological space is a pair (H, τ) where τ is a *topology*. This can be constructed from open sets which are defined in terms of a distance function. The *generated σ-algebra* is then the smallest σ-algebra that contains these open sets. An example for a topological space is $(\mathbb{R}, O)$, where O contains all open intervals of $\mathbb{R}$. Then the generated σ-algebra denoted by $\sigma(O)$ is the well–known Borel σ-algebra.

Lemma 2.4 (Continuous functions are measurable [79, Theorem 1.88])**.** Let $Z : H_1 \to H_2$ be a continuous function w.r.t. the topological spaces (H_1, τ_1) and (H_2, τ_2). Then Z is measurable w.r.t. the measurable spaces $(H_1, \mathcal{G}_1)$ and $(H_2, \mathcal{G}_2)$, where $\mathcal{G}_1 := \sigma(\tau_1)$ and $\mathcal{G}_2 := \sigma(\tau_2)$ denotes the generated σ-algebra.

Proof. The main idea of the proof by Klenke [79, Theorem 1.88] is that the preimage of an open set is open for the continuous function Z. □

We often examine two or more random variables and their relationship. A pair of random variables (Z, Y) is independent if we are allowed to examine Z and Y separately. Informally, this means that we do not gain any information regarding the value of Z even if we know the value of Y and vice versa. We make this statement precise for a finite set of random variables.

Definition 2.5 (Independence of random variables [79, Remark 2.15])**.** Let $(Z_1, \dots, Z_L)$ be a random vector with associated probability space

$$(H_1 \times \cdots \times H_L, \mathcal{G}_1 \times \cdots \times \mathcal{G}_L, \mathbb{P}((Z_1, \dots, Z_L) \in \cdot)).$$

We denote the *marginal probability measure* of Z_ℓ with $\mathbb{P}_\ell$ for all $\ell \in \{1, \dots, L\}$. Then the random variables $Z_1, \dots, Z_L$ are called *independent* if the probability measure $\mathbb{P}$ factorizes such that for all $G_1 \in \mathcal{G}_1, \dots, G_L \in \mathcal{G}_L$

$$\mathbb{P}((Z_1, \dots, Z_L) \in G_1 \times \cdots \times G_L) = \prod_{\ell=1}^{L} \mathbb{P}_\ell(Z_\ell \in G_\ell). \qquad \diamond$$

Simulation based techniques often require multiple realizations or samples of a random variable Z. The idea is to extract information by looking at independent copies of Z which are evaluated for some event $\omega \in \Omega$.

Definition 2.6 (Independent identically distributed samples)**.** The random variables $Z^1, \dots, Z^m$ are *independent identically distributed* (i.i.d.) if $Z^1, \dots, Z^m$ are independent and $Z^\ell \sim \mathbb{P}$ for all $\ell \in \{1, \dots, m\}$. We always assume that random variables with different superscripts are i.i.d.. For $\omega \in \Omega$ we call $Z(\omega)$ a *sample* or *realization* of the random variable Z. By slight abuse of notation we often drop the ω and denote *i.i.d. samples* of Z with $Z^1 := Z^1(\omega), \dots, Z^m := Z^m(\omega)$. $\diamond$

The computation of a sample often incurs a random computational cost. This definition is rather vague, since the exact cost may be the actual time a computer needs to compute a sample. We may also define the cost as degrees of freedom or number of operations if computing the sample requires us to solve a linear system.

Definition 2.7 (Expected cost for a random variable)**.** The cost function $\mathbb{W}$ maps a random variable Z to a non–negative real number

$$\mathbb{W} : \{Z : \Omega \to H \mid Z \text{ measurable } \} \to \mathbb{R}_{\geq 0}. \qquad \diamond$$

The value of $\mathbb{W}[Z]$ is interpreted as expected cost to compute a sample of Z. For random variables $Z_1, \dots, Z_L$ we abbreviate

$$w_\ell := \mathbb{W}[Z_\ell] \quad \text{for all } \ell \in \{1, \dots, L\}.$$

We frequently make statements about random variables that are certain and occur with probability one. We formally define this and the equivalent formulation for a general measure space.

Definition 2.8 (Almost all, $\mathbb{P}$–almost surely)**.** Let $(\Omega, \mathcal{F}, \nu)$ be a measure space and $f : H \to \{0, 1\}$ be a measurable function. We say that the property f holds for *ν–almost all* $\omega \in \Omega$ if

$$\nu(f^{-1}(\{0\})) = 0.$$

If $\nu := \mathbb{P}$ is a probability measure we say that f holds *$\mathbb{P}$–almost surely*, which we sometimes abbreviate with $\mathbb{P}$–a.s.. We often drop $\mathbb{P}$ from the notation. ◇

There is little value in specifying the measurable space $(\Omega, \mathcal{F})$, since this is often given implicitly in terms of a probability measure $\mathbb{P}$, a probability density function, a random variable or a cumulative distribution function. Therefore, we never attempt to describe both Ω and $\mathcal{F}$ in this thesis. We however, always specify $\mathbb{P}$ or a corresponding random variable Z.

Moments. For notational purposes and throughout the rest of this chapter we assume that Z and Y assume values in the Hilbert space H unless stated otherwise. It is often helpful to summarize or compress a function or random variable Z into a single value. We achieve this if we integrate out the domain of Z raised to some power p. This operation is well defined for measurable functions if the function is p-integrable. We formulate this for a general measure space, where we denote the *scalar product* of a Hilbert space H with $(\cdot, \cdot)_H$ and the *induced norm* with $\|\cdot\|_H$.

Definition 2.9 (Lebesgue space L^p)**.** Let $(\Omega, \mathcal{F}, \nu)$ be a measure space and $p \in [1, +\infty]$. The *Lebesque space* $L^p(\Omega, H, \nu)$ is the space of measurable functions whose *p–th moment* is bounded

$$L^p(\Omega, H, \nu) := \big\{ Z : \Omega \to H \,|\, Z \text{ is } \nu\text{–measurable and } \|Z\|_{L^p(\Omega,H,\nu)} < +\infty \big\},$$

where the norm $\|\cdot\|_{L^p(\Omega,H,\nu)}$ for $p \in [1, +\infty)$ is defined such that

$$\|Z\|^p_{L^p(\Omega,H,\nu)} := \int_\Omega \|Z(\omega)\|^p_H d\nu(\omega).$$

For the special case $p = +\infty$ the norm is defined as

$$\|Z\|_{L^\infty(\Omega,H,\nu)} := \sup\{c \in \mathbb{R} \,|\, \nu(\|Z\|_H \leq c) > 0\}.$$

If the meaning of the space is clear from the context we use the abbreviation

$$L^p := L^p(\Omega) := L^p(\Omega, H) := L^p(\Omega, H, \nu).$$ ◇

It is well known that the space L^p is a *Banach space* if we identify functions that are equal up to a set of measure zero. Furthermore, for $p = 2$ the space L^2 is a Hilbert space that inherits the inner product structure from H. We formally define this inner product such that for all $Z, Y \in L^2$

$$(Z, Y)_{L^2} := \int_\Omega (Z(\omega), Y(\omega))_H d\nu(\omega). \tag{2.1}$$

The Cauchy–Schwarz inequality shows that (2.1) is well defined. This inequality is a special case of Hölder's inequality and we precisely state both now.

Lemma 2.10 (Hölder's inequality, Cauchy–Schwarz inequality)**.** Let $p, q \in [1, +\infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, $Z \in L^p$ and $Y \in L^q$. Then *Hölder's inequality* holds

$$\| (Z, Y)_H \|_{L^1} \leq \|Z\|_{L^p} \|Y\|_{L^q}.$$

The special case with $p = q = 2$ is the *Cauchy–Schwarz inequality*

$$\| (Z, Y)_H \|_{L^1} \leq \|Z\|_{L^2} \|Y\|_{L^2}.$$

Proof. See Klenke [79, Chapter 7]. □

A straightforward consequence of Hölder's inequality and $\mathbb{P}(\Omega) = 1$ is that for all $p \in [1, +\infty]$ the random variable $Z \in L^p(\Omega, H, \mathbb{P})$ is also an element of the space $L^q(\Omega, H, \mathbb{P})$ for all $q \in [1, p]$. This implication is in general not true if we replace $\mathbb{P}$ with an arbitrary measure ν.

Expectation, Variance, Covariance and Correlation. We proceed to define specific integrals concerning random variables. The expectation or mean of a random variable describes its average value. We use the variance to describe the average squared deviation from the mean. Both values are basic properties of random variables and exist if the first respectively second moment is finite.

Definition 2.11 (Expectation, Variance)**.** For $Z \in L^1$ we define the *expectation* or *mean*

$$\mathbb{E}[Z] := \int_\Omega Z(\omega) d\mathbb{P}(\omega).$$

This definition has to be understood as Bochner integral if Z is not real–valued. If in addition $Z \in L^2$ we define the *variance*

$$\mathbb{V}[Z] := \mathbb{E}\big[\|Z - \mathbb{E}[Z]\|_H^2\big] = \mathbb{E}\big[\|Z\|_H^2\big] - \|\mathbb{E}[Z]\|_H^2. \tag{2.2}$$

◇

For real–valued Z the variance (2.2) coincides with the usual definition. For random variables $Z_1, \ldots, Z_L \in L^2$ we abbreviate the mean and variance

$$\begin{aligned} \mu_\ell &:= \mathbb{E}[Z_\ell] && \text{for all } \ell \in \{1, \ldots, L\}, \\ \sigma_\ell^2 &:= \mathbb{V}[Z_\ell] && \text{for all } \ell \in \{1, \ldots, L\}. \end{aligned}$$

The mean is the constant in H that best approximates the random variable Z. It is the unique solution of the minimization problem

$$\min_{\mu \in H} \|Z - \mu\|_{L^2}^2 = \mathbb{E}\big[\|Z - \mu\|_H^2\big].$$

The value of the cost function at the minimizer $\mu = \mathbb{E}[Z]$ is the variance and describes the approximation error. The variance is zero $\mathbb{V}[Z] = 0$ if and only if Z is almost surely constant with $Z = \mathbb{E}[Z]$. In all other cases the variance is positive. We are allowed to pull out constants from the variance by squaring them, that is for all $\beta \in \mathbb{R}$

$$\mathbb{V}[\beta Z] = \beta^2 \mathbb{V}[Z].$$

The expectation is a linear operator since for all $Z_1, \dots, Z_L \in L^1$ and vectors $\beta \in \mathbb{R}^L$

$$\mathbb{E}\left[\sum_{\ell=1}^{L} \beta_\ell Z_\ell\right] = \sum_{\ell=1}^{L} \beta_\ell \mathbb{E}[Z_\ell].$$

The mean and variance are concerned with a single random variable Z and both give a single value. We now describe relationships between two random variables with the help of the covariance and correlation.

Definition 2.12 (Covariance, Correlation)**.** For real–valued random variables $Z, Y \in L^2$ we define the *covariance*

$$\mathbb{C}\mathrm{ov}[Z, Y] := \mathbb{E}[(Z - \mathbb{E}[Z])(Y - \mathbb{E}[Y])] = \mathbb{E}[ZY] - \mathbb{E}[Z]\mathbb{E}[Y].$$

If in addition $\mathbb{V}[Z], \mathbb{V}[Y] > 0$ we define the *correlation* or *correlation coefficient*

$$\mathbb{C}\mathrm{orr}[Z, Y] := \frac{\mathbb{C}\mathrm{ov}[Z, Y]}{(\mathbb{V}[Y]\mathbb{V}[Z])^{1/2}}.$$

The random variables Z and Y are *uncorrelated* if $\mathbb{C}\mathrm{ov}[Z, Y] = 0$. ◇

We extend the previous definition to multiple random variables. This allows us to place multiple covariance and correlation values into a vector or matrix.

Definition 2.13 (Covariance and Correlation matrix)**.** For vectors of real–valued random variables $Z := (Z_1, \dots, Z_L)^T \in L^2$ and $Y := (Y_1, \dots, Y_N)^T \in L^2$ we define the *covariance matrix*

$$\mathbb{C}\mathrm{ov}[Z, Y] := \mathbb{E}\big[(Z - \mathbb{E}[Z])(Y - \mathbb{E}[Y])^T\big] \in \mathbb{R}^{L \times N}.$$

For $\mathbb{V}[Z_1], \dots, \mathbb{V}[Z_L], \mathbb{V}[Y_1], \dots, \mathbb{V}[Y_N] > 0$ we define the *correlation matrix* as the correlation between entries of Z and Y

$$\mathbb{C}\mathrm{orr}[Z, Y] \in \mathbb{R}^{L \times N},\ \mathbb{C}\mathrm{orr}[Z, Y]_{\ell,n} := \mathbb{C}\mathrm{orr}[Z_\ell, Y_n] \quad \text{for all } \ell \in \{1, \dots, L\}, n \in \{1, \dots, N\}.$$

◇

For real–valued random variables $Z_1, \dots, Z_L \in L^2$ we abbreviate

$$C := \mathbb{C}\mathrm{ov}\left[\begin{pmatrix} Z_1 \\ \vdots \\ Z_L \end{pmatrix}, \begin{pmatrix} Z_1 \\ \vdots \\ Z_L \end{pmatrix}\right] \in \mathbb{R}^{L \times L}, \quad \mathrm{P} := (\rho_{ij})_{i,j=1}^{L} := \mathbb{C}\mathrm{orr}\left[\begin{pmatrix} Z_1 \\ \vdots \\ Z_L \end{pmatrix}, \begin{pmatrix} Z_1 \\ \vdots \\ Z_L \end{pmatrix}\right] \in \mathbb{R}^{L \times L}. \tag{2.3}$$

We generalize the covariance to infinite dimensional Hilbert spaces. The basic idea is to reduce the Hilbert space valued random variable Z to a single value in $\mathbb{R}$ by testing it with a linear functional in the dual space H^*. We identify this space with H due to the Riesz–representation theorem [79, Theorem 7.26].

Definition 2.14 (Covariance operator, Correlation operator)**.** Let H_1, H_2 be real Hilbert spaces. For $Z \in L^2(\Omega, H_1)$ and $Y \in L^2(\Omega, H_2)$ we define the *covariance operator*

$$\mathbb{C}\mathrm{ov}[Z, Y] : H_1 \times H_2 \to \mathbb{R}, \quad \mathbb{C}\mathrm{ov}[Z, Y](z, y) := \mathbb{E}[(z, Z - \mathbb{E}[Z])_{H_1}(y, Y - \mathbb{E}[Y])_{H_2}].$$

We define the *correlation operator* accordingly

$$\mathbb{C}\mathrm{orr}[Z, Y] : H_1 \times H_2 \to \mathbb{R}, \quad \mathbb{C}\mathrm{orr}[Z, Y](z, y) := \frac{\mathbb{C}\mathrm{ov}[Z, Y](z, y)}{(\mathbb{C}\mathrm{ov}[Z, Z](z, z)\mathbb{C}\mathrm{ov}[Y, Y](y, y))^{1/2}},$$

whenever the quotient is not equal to zero. ◇

The covariance operator is a generalization of the covariance matrix where the Hilbert spaces are $H_1 := \mathbb{R}^L$ and $H_2 := \mathbb{R}^N$ with the Euclidean inner product. We obtain the entries of the covariance matrix if we test with the unit vectors $z := e_\ell$ and $y := e_n$

$$\begin{aligned}\mathbb{C}\mathrm{ov}[Z,Y](z,y) &= \mathbb{E}[(e_\ell, Z - \mathbb{E}[Z])_{\mathbb{R}^L}(e_n, Y - \mathbb{E}[Y])_{\mathbb{R}^N}] \\ &= \mathbb{E}[(Z_\ell - \mathbb{E}[Z_\ell])(Y_n - \mathbb{E}[Y_n])] \\ &= \mathbb{C}\mathrm{ov}[Z,Y]_{\ell,n}.\end{aligned}$$

Let us now verify that the covariance is actually well defined for $Z, Y \in L^2$. We fix $z \in H_1$, $y \in H_2$, apply the Cauchy–Schwarz inequality twice and use the linearity of the expectation

$$\begin{aligned}\mathbb{C}\mathrm{ov}[Z,Y](z,y)^2 &\leq \mathbb{E}\big[(z, Z - \mathbb{E}[Z])_{H_1}^2\big]\mathbb{E}\big[(y, Y - \mathbb{E}[Y])_{H_2}^2\big] \\ &\leq \mathbb{E}\big[\|z\|_{H_1}^2 \|Z - \mathbb{E}[Z]\|_{H_1}^2\big]\mathbb{E}\big[\|y\|_{H_2}^2 \|Y - \mathbb{E}[Y]\|_{H_2}^2\big] \\ &= \|z\|_{H_1}^2 \|y\|_{H_2}^2 \mathbb{E}\big[\|Z - \mathbb{E}[Z]\|_{H_1}^2\big]\mathbb{E}\big[\|Y - \mathbb{E}[Y]\|_{H_2}^2\big].\end{aligned}$$

The last term is bounded for $Z, Y \in L^2$. We summarize some well–known properties of the covariance.

Lemma 2.15 (Properties of the covariance)**.** For $Z, Z_1, \ldots, Z_L \in L^2(\Omega, H_1)$ and $Y \in L^2(\Omega, H_2)$ the covariance operator $\mathbb{C}\mathrm{ov}$ is

- symmetric: $\mathbb{C}\mathrm{ov}[Z,Y](z,y) = \mathbb{C}\mathrm{ov}[Y,Z](y,z)$ for all $z \in H_1, y \in H_2$.
- bilinear: $\mathbb{C}\mathrm{ov}\Big[\sum_{\ell=1}^L \beta_\ell Z_\ell, Y\Big] = \sum_{\ell=1}^L \beta_\ell \mathbb{C}\mathrm{ov}[Z_\ell, Y]$ for all $\beta \in \mathbb{R}^L$.
- positive semi–definite: $\mathbb{C}\mathrm{ov}[Z,Z](z,z) \geq 0$ for all $z \in H$.
- equal to the variance if Z is real–valued: $\mathbb{C}\mathrm{ov}[Z,Z] = \mathbb{V}[Z]$.

Proof. The properties follow directly from the definition of the covariance. □

The correlation $\mathbb{C}\mathrm{orr}$ inherits its properties from the covariance. The Cauchy–Schwarz inequality can be used to show that the correlations operator takes values between -1 and 1. Formally, for all $z \in H_1$ and $y \in H_2$ where the correlation is well defined

$$\mathbb{C}\mathrm{orr}[Z,Y](z,y) \in [-1,1].$$

The covariance matrix $C \in \mathbb{R}^{L\times L}$ is always positive semi–definite. We now show that if the entries of $Z - \mathbb{E}[Z]$ with $Z = (Z_1, \ldots, Z_L)^T$ are linearly independent, then $C = \mathbb{C}\mathrm{ov}[Z,Z]$ is positive definite and thus invertible.

Lemma 2.16 (Positive definiteness of the covariance)**.** For $Z \in L^2$ the following statements are equivalent

- $\mathbb{C}\mathrm{ov}[Z,Z](z,z) > 0$ for all $z \in H \setminus \{0\}$.
- $(z, Z - \mathbb{E}[Z])_H \neq 0$ $\mathbb{P}$–almost surely for all $z \in H \setminus \{0\}$.

For $Z := (Z_1, \ldots, Z_L)^T$ with real–valued $Z_1, \ldots, Z_L$ this equivalence reads

- $\beta^T \mathbb{C}\mathrm{ov}[Z,Z]\beta > 0$ for all $\beta \in \mathbb{R}^L \setminus \{0\}$.

- $Z_1 - \mathbb{E}[Z_1], \dots, Z_L - \mathbb{E}[Z_L]$ are linearly independent.

Furthermore, if $\beta^T \mathbb{C}\mathrm{ov}[Z, Z]\beta = 0$ for some $\beta \in \mathbb{R}^L$ then $\mathbb{P}$–almost surely

$$\beta^T (Z - \mathbb{E}[Z]) = 0. \tag{2.4}$$

Proof. We deduce the claim by directly looking at the definition of the covariance

$$\mathbb{C}\mathrm{ov}[Z, Z](z, z) = \mathbb{E}\big[(z, Z - \mathbb{E}[Z])_H^2\big]. \qquad \square$$

The factorization of the probability measure $\mathbb{P}$ for independent random variables shows that their covariance is zero. We summarize this and further properties in the next lemma.

Lemma 2.17 (Properties of independent random variables [79, Theorem 5.4])**.** Let $Z, Y \in L^1$ be independent random variables. Then the expectation of the product is equal to the product of the expectations

$$\mathbb{E}[(Z, Y)_H] = (\mathbb{E}[Z], \mathbb{E}[Y])_H.$$

For $Z, Y \in L^2$ the random variables Z and Y are uncorrelated $\mathbb{C}\mathrm{ov}[Z, Y] = 0$. $\diamond$

The previous lemma shows that independence of Z, Y implies that Z, Y are uncorrelated, however the converse is in general not true. We now state an important computational rule for the variance of sums of random variables.

Lemma 2.18 (Variance of sums [79, Theorem 5.7])**.** Let $Z_1, \dots, Z_L \in L^2$ be real–valued random variables. Then the variance of the sum satisfies

$$\mathbb{V}\left[\sum_{\ell=1}^{L} Z_\ell\right] = \sum_{\ell,j=1}^{L} \mathbb{C}\mathrm{ov}[Z_\ell, Z_j] = \sum_{\ell=1}^{L} \mathbb{V}[Z_\ell] + \sum_{\substack{\ell,j=1 \\ \ell \neq j}}^{L} \mathbb{C}\mathrm{ov}[Z_\ell, Z_j].$$

If $Z_1, \dots, Z_L$ are pairwise uncorrelated then the covariance terms are equal to zero

$$\mathbb{V}\left[\sum_{\ell=1}^{L} Z_\ell\right] = \sum_{\ell=1}^{L} \mathbb{V}[Z_\ell].$$

Proof. We use the bilinearity of the covariance and $\mathbb{C}\mathrm{ov}[Z_\ell, Z_j] = 0$ for uncorrelated Z_ℓ, Z_j. $\square$

Convergence of random variables. There are different convergence types for random variables. In this thesis we distinguish between almost sure convergence and convergence in the Lebesgue space L^p.

Definition 2.19 (Almost sure convergence [79, Definition 6.2])**.** Let $(Z_n)_{n=1}^{\infty}$ be a sequence of random variables. We say that $(Z_n)_{n=1}^{\infty}$ *converges almost surely* to the random variable Z if

$$\mathbb{P}\left(\lim_{n \to +\infty} Z_n = Z\right) = 1. \qquad \diamond$$

Definition 2.20 (Convergence in L^p [79, Definition 7.2])**.** Let $(Z_n)_{n=1}^{\infty} \subseteq L^p$ and $Z \in L^p$. Then $(Z_n)_{n=1}^{\infty}$ converges to Z in L^p if

$$\lim_{n \to +\infty} \|Z_n - Z\|_{L^p} = 0. \qquad \diamond$$

Useful inequalities. The probability that a random variable deviates from its mean is bounded by its variance. This allows us to estimate the probability that the realizations of a random variable remain within a certain distance from its mean.

Theorem 2.21 (Markov inequality, Chebyshev inequality, [79, Theorem 5.11])**.**
Let Z be a random variable and $f : [0, +\infty) \to [0, +\infty)$ a monotonically increasing function. Then for all $\varepsilon > 0$ *Markov's inequality* holds

$$\mathbb{P}(\|Z\|_H \geq \varepsilon) \leq \frac{\mathbb{E}[f(\|Z\|_H)]}{f(\varepsilon)}$$

provided that the right–hand side is well defined. For $Z \in L^2$ the *Chebyshev inequality* holds

$$\mathbb{P}(\|Z - \mathbb{E}[Z]\|_H \geq \varepsilon) \leq \frac{\mathbb{V}[Z]}{\varepsilon^2}. \qquad \diamond$$

We state Jensen's inequality, which allows us to exchange the expectation and a convex function φ at the cost of introducing an inequality.

Lemma 2.22 (Jensen's inequality [79, Theorem 7.11])**.** Let $I := (a, b) \subseteq \mathbb{R}$ be an open interval, $Z \in L^1(\Omega, I)$ and the function $\varphi : I \to \mathbb{R}$ convex. Then *Jensen's inequality* holds

$$\varphi(\mathbb{E}[Z]) \leq \mathbb{E}[\varphi(Z)],$$

provided that the right–hand side is well defined.

Proof. A formal proof is given in [79, Theorem 7.11]. We only remark that convex functions are continuous and thus measurable, hence $\varphi(Z)$ is a random variable. $\square$

We finish this section with the elementary Young inequality, which we use to estimate the expectation of a product of real–valued random variables such that $2\mathbb{E}[ZY] \leq \mathbb{E}[Z^2] + \mathbb{E}[Y^2]$.

Lemma 2.23 (Young inequality [79, Lemma 7.15])**.** For $p, q \in (1, +\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ and real numbers $z, y \in [0, +\infty)$ *Young's inequality* holds

$$zy \leq \frac{z^p}{p} + \frac{y^q}{q}.$$

In particular, for $p = q = 2$ the inequality holds for arbitrary $z, y \in \mathbb{R}$. $\diamond$

2.2 Karhunen–Loève expansion

Construction of Gaussian random fields. The *Karhunen–Loève expansion* (KLE) is a powerful tool to generate random variables with values in an infinite dimensional, separable Hilbert space. The main idea is to randomize the coefficients of a Fourier series in a suitable way such that the series converges almost surely. A study of orthogonal expansions of random variables is available in [1, Section 3]. First, we provide conditions for the special case of real–valued random variables.

Lemma 2.24 (Khinchin and Kolmogorov [77, Lemma 3.16])**.** Let $(\xi_n)_{n=1}^{\infty}$ be real–valued independent random variables with $\mathbb{E}[\xi_n] = 0$ for all $n \in \mathbb{N}$ such that their variance is summable

$$\sum_{n=1}^{\infty} \mathbb{V}[\xi_n] < +\infty.$$

Then the series $\sum_{n=1}^{\infty} \xi_n$ converges almost surely. ◇

We require the previous lemma to ensure that the KLE is well defined.

Definition 2.25 (Infinite dimensional Gaussian random field)**.** Let H be an infinite dimensional, separable, real Hilbert space. Furthermore, assume the following:

1. The eigenfunctions $(\psi_n)_{n=1}^{\infty}$ form a complete orthonormal basis of H,
2. The random variables $(\xi_n)_{n=1}^{\infty}$ are i.i.d. standard normals with $\xi_n \in N(0,1)$ for all $n \in \mathbb{N}$,
3. The eigenvalues $(\lambda_n)_{n=1}^{\infty}$ are non–negative values $\lambda_n \geq 0$ for all $n \in \mathbb{N}$ and the sequence is summable

$$\sum_{n=1}^{\infty} \lambda_n < +\infty. \tag{2.5}$$

We define the KLE a as series

$$a := \sum_{n=1}^{\infty} \sqrt{\lambda_n} \xi_n \psi_n.$$

We have $a \in L^2$ and define $\mathcal{C} := \mathbb{C}\mathrm{ov}[a,a]$, which we abbreviate with $a \sim N(0,\mathcal{C})$. ◇

We use Parseval's identity and apply Lemma 2.24 in combination with (2.5) to show that the norm of a is almost surely bounded

$$\|a\|_H^2 = \sum_{n=1}^{\infty} \lambda_n \xi_n^2 = \sum_{n=1}^{\infty} \lambda_n (\xi_n^2 - 1) + \sum_{n=1}^{\infty} \lambda_n < +\infty.$$

We thus conclude that $a \in H$ almost surely. We use the monotone convergence theorem [77, Theorem 1.19] to exchange the mean and summation

$$\mathbb{E}\big[\|a\|_H^2\big] = \mathbb{E}\left[\sum_{n=1}^{\infty} \lambda_n \xi_n^2\right] = \sum_{n=1}^{\infty} \mathbb{E}\big[\lambda_n \xi_n^2\big] = \sum_{n=1}^{\infty} \lambda_n < +\infty,$$

which shows that $a \in L^2$. The dominated convergence theorem [77, Theorem 1.21] now shows that $\mathbb{E}[a] = 0$.

We clarify why we call λ_n eigenvalues and ψ_n the eigenfunctions of $\mathcal{C}(\psi_n, \cdot)$. The functions $(\psi_n)_{n=1}^{\infty}$ form a complete orthonormal basis and thus

$$\mathcal{C}(\psi_n, \cdot) = \sum_{k,m=1}^{\infty} \sqrt{\lambda_k}\sqrt{\lambda_m}\mathbb{E}[\xi_k\xi_m](\psi_n, \psi_k)_H(\cdot, \psi_m)_H = \sum_{m=1}^{\infty} \sqrt{\lambda_n}\sqrt{\lambda_m}\mathbb{E}[\xi_n\xi_m](\cdot, \psi_m)_H.$$

We use the independence of ξ_n and ξ_m for $n \neq m$, $\mathbb{E}[\xi_m] = 0$ and $\mathbb{E}[\xi_m^2] = 1$ to conclude

$$\sum_{m=1}^{\infty} \sqrt{\lambda_n}\sqrt{\lambda_m}\mathbb{E}[\xi_n\xi_m](\cdot, \psi_m)_H = \lambda_n(\cdot, \psi_n)_H.$$

We interpret $\mathcal{C}(\psi_n, \cdot)$ as an element of H with the help of the Riesz–representation theorem and conclude that ψ_n is an eigenfunction of $\mathcal{C}$ with eigenvalue λ_n

$$\mathcal{C}(\psi_n) = \lambda_n\psi_n.$$

The covariance operator is diagonal $\mathcal{C}(\psi_n, \psi_j) = \lambda_n\delta_{nj}$, where δ_{nj} is the Kronecker delta

$$\delta_{nj} := \begin{cases} 1, & \text{if } n = j, \\ 0, & \text{if } n \neq j. \end{cases}$$

We are interested in random fields and thus ψ_n are functions. For an event ω and $x \in D$ the KLE with arguments has the following form

$$a(x, \omega) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\xi_n(\omega)\psi_n(x).$$

We want to construct a random field such that values at close points are highly correlated. For points $x, y \in D$ we require that

$$\mathbb{C}\mathrm{ov}[a(x), a(y)] = k(\|x - y\|),$$

where k is a stationary covariance kernel and $\|\cdot\|$ a suitable norm. We call a covariance kernel or random field *stationary* if the covariance $\mathbb{C}\mathrm{ov}[a(x), a(y)]$ depends only on the distance between x and y. We work with the commonly used Whittle–Matérn covariance kernel. Practical applications of covariance kernels are Gaussian processes regression or kriging in machine learning [115], where the kernel models the similarity between the datapoints. Another application is spatial descriptions in geostatistics [33, 140].

Definition 2.26 (Whittle–Matérn covariance kernel [115, Chapter 4])**.**
The *Whittle–Matérn covariance kernel* k has three parameters, the *variance* $\sigma^2 > 0$, the *smoothness* $\nu > 0$ and the *correlation length* $\ell > 0$. We write down the kernel for different values of ν for a given distance $d \geq 0$

- $\nu = 1/2$: $\quad k(d) := \sigma^2 \exp(-d/\ell)$,
- $\nu = 3/2$: $\quad k(d) := \sigma^2(1 + \sqrt{3}d/\ell)\exp(-\sqrt{3}d/\ell)$,
- $\nu = 5/2$: $\quad k(d) := \sigma^2\big(1 + \sqrt{5}d/\ell + 5d^2/(3\ell^2)\big)\exp(-\sqrt{5}d/\ell)$. ◇

The Whittle–Matérn covariance kernel now defines the KLE of a Gaussian.

Theorem 2.27 (Gaussian from Whittle–Matérn covariance kernel [115, Chapter 4])**.** Let $D \subseteq \mathbb{R}^d$ be a bounded domain. Then there exists $a \sim N(0, \mathcal{C})$ with values in $C(\overline{D})$ such that for all $x, y \in D$

$$\mathbb{C}\mathrm{ov}[a(x), a(y)] = k(\|x - y\|),$$

where k is the Whittle–Matérn covariance kernel. We express the random field a as KLE

$$a := \sum_{n=1}^{\infty} \sqrt{\lambda_n} \psi_n \xi_n.$$

We restrict ourselves to the random field $a \sim N(0, \mathcal{C})$ where $\mathcal{C}$ is induced by a Whittle–Matérn covariance kernel k. Other stationary random fields are examined by Adler [1, Section 5] and a general theory on the existence of Gaussian random fields in infinite dimensions is given by Bogachev [12].

Practical implementation. The KLE for a is a series

$$a = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \xi_n \psi_n,$$

which we truncate to obtain a finite sum that we evaluate numerically. We define the *truncated KLE*

$$a_N := \sum_{n=1}^{N} \sqrt{\lambda_n} \xi_n \psi_n.$$

The expected value of the error for this truncation in the H-norm is equal to the sum of the missing eigenvalues.

Lemma 2.28 (KLE truncation error)**.** The expected truncation error for a KLE is

$$\mathbb{E}\big[\|a_N - a\|_H^2\big] = \sum_{n=N+1}^{\infty} \lambda_n. \tag{2.6}$$

This error converges to zero for $N \to +\infty$.

Proof. The error (2.6) follows from the Parseval identity and converges to zero since the eigenvalues are summable by assumption (2.5). □

For a fixed truncation index N we minimize the truncation error if we keep the N largest eigenvalues. We thus sort the eigenvalues in descending order

$$\lambda_1 \geq \lambda_2 \geq \dots.$$

The eigenpairs (λ_n, ψ_n) for the Whittle–Matérn covariance typically cannot be computed analytically. However, for arbitrary $v \in L^2(D)$ the eigenpairs are related to the kernel k in the following way

$$\begin{aligned}
\lambda_n (\psi_n, v)_{L^2(D)} &= \mathbb{C}\mathrm{ov}[a, a](\psi_n, v) \\
&= \int_D \int_D \psi_n(x) \mathbb{E}[a(x)a(y)] v(y) dx dy \\
&= \int_D \int_D \psi_n(x) k(\|x - y\|) v(y) dx dy \\
&= ((\psi_n, k(\|\cdot - \cdot\|))_{L^2(D)}, v)_{L^2(D)}.
\end{aligned}$$

This expression holds for all $v \in L^2(D)$ and thus for almost all $y \in D$

$$(k(\|\cdot - y\|), \psi_n)_{L^2(D)} = \lambda_n \psi_n(y). \tag{2.7}$$

The collocation approximation solves a numeric approximation of (2.7), which is equivalent to a matrix eigenvalue problem.

Definition 2.29 (Collocation approximation [92, Section 7.4])**.** Let $w_1, \dots, w_M \in \mathbb{R}$ be quadrature weights and $x_1, \dots, x_M \in D$ collocation points such that

$$\sum_{j=1}^{M} w_j k(\|x_j - x_m\|)\psi(x_j) = \lambda \psi(x_m) \quad \text{for all } m \in \{1, \dots, M\}.$$

The *collocation approximation* requires us to solve the following eigenvalue problem

$$\begin{aligned} &Av = \lambda v, \\ &A \in \mathbb{R}^{M \times M}, && A_{jm} := w_j k(\|x_j - x_m\|), \\ &v \in \mathbb{R}^M, && v_m := \psi(x_m). \end{aligned}$$ ◇

The truncated KLE allows us to compute only the N largest eigenvalues and the corresponding eigenvectors. This problem is known as *(generalized) Hermitian eigenvalue problem*, which is a well–studied problem in linear algebra, see for example [61, Chapter 8]. The evaluation of a_N at a point $x \in D$ is then approximately given by

$$a_N(x) \approx \sum_{n=1}^{N} \sqrt{\lambda_n} \xi_n v^n[x],$$

where $v^n[x] \approx \psi_n(x)$ is a suitable interpolation that uses the entries in the n–th eigenvector v^n. This procedure allows us to draw samples of a_N which are close to a.
We give some examples for a Gaussian mean zero random field with Whittle–Matérn covariance in Figure 2.1 for $D := [0, 1]$. Figure 2.2 contains samples with similar parameters for $D := [0, 1]^2$. The parameter σ^2 for the Whittle–Matérn kernel in Definition 2.26 controls the variance or magnitude of the random field and is fixed $\sigma^2 := 1$ for all examples. The correlation length ℓ controls how fast the kernel decays. Large values of ℓ ensure that $a(x) \approx a(y)$ even if x is not close to y. Small values of ℓ ensure that $a(x)$ and $a(y)$ are almost uncorrelated even if x is close to y. Increasing the smoothness parameter ν leads to smoother realizations.
We remark that the collocation approximation is not the only approach to sample from a random field. The Galerkin approach [92, Section 7.4] approximates the eigenfunctions ψ in a finite dimensional subspace of $L^2(D)$. The algorithm of Saibaba [124] computes a random solution for the eigenvalue problem (2.7). For stationary random fields there are circulant embedding methods to sample efficiently, see [47] or [92, Section 7.2]. We remark that this is not an exhaustive list of methods to sample from a Gaussian with values in an infinite dimensional Hilbert space.

2.3 Elliptic partial differential equation and discretization

Models with uncertain inputs. In this section we write down mathematical models that we use for our numerical experiments. Let H_1, H_2 be two measurable spaces. A

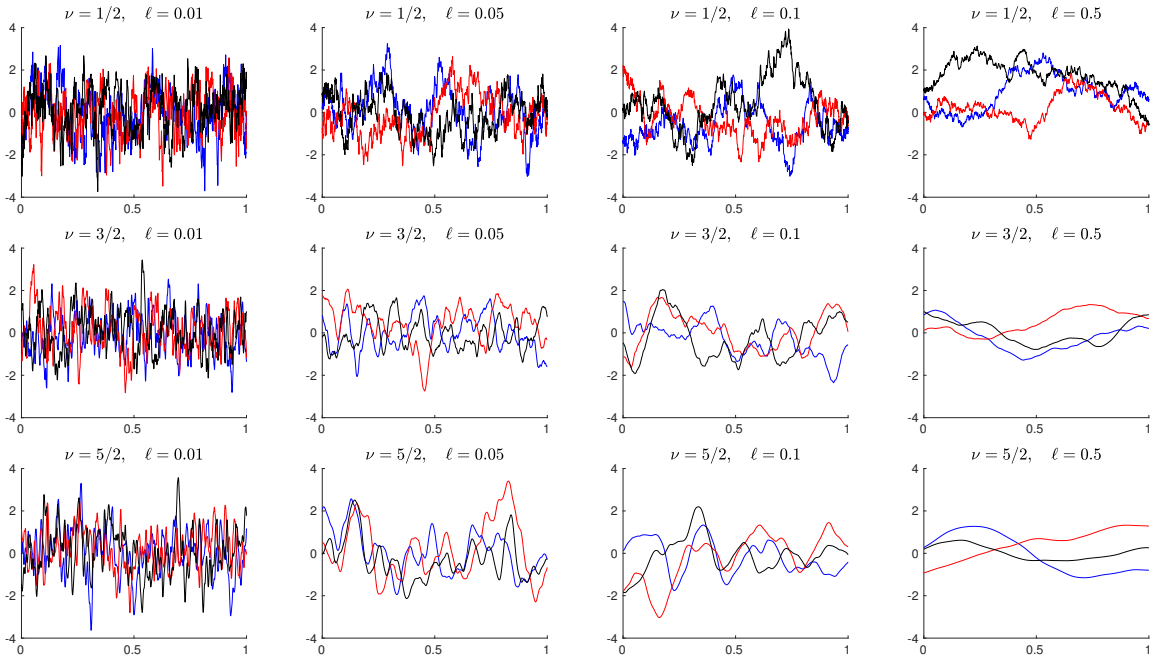

Figure 2.1: Samples of a Gaussian mean zero random field with Whittle–Matérn covariance in $D = [0, 1]$ for different values of the smoothness $\nu \in \{1/2, 3/2, 5/2\}$, correlation length $\ell \in \{0.01, 0.05, 0.1, 0.5\}$ and fixed variance $\sigma^2 = 1$. Each plot shows three independent realizations (red, blue and black). We generated the plots by a truncated KLE after $N = 1000$ terms, $M = 1000$ collocation points and each random field is evaluated at 10^4 points.

model consists of an input parameter $a \in H_1$ and a continuous function $f : H_1 \to H_2$ which maps an input to the *Quantity of Interest* (QoI)

$$Z := f(a) \in H_2.$$

We create a random model if we replace the deterministic input a with a random variable $a : \Omega \to H_1$ such that the QoI is also a random variable

$$Z := f \circ a : \Omega \to H_2.$$

Our goal is to collect information of Z like the mean, variance and so forth. The justification for such an approach is that in practice the input quantities are often either unknown or known only up to a certain accuracy. Instead of a deterministic input we allow a random input which assumes more values where each one is weighted according to the distribution $\mathbb{P}$ of a. The gain is that we look at more than a single configuration of the model and by examining the random Z we better understand how uncertainties in the input a propagate to the QoI.

Poisson's equation. We fix some notation for differential operators. For an open set $D \subseteq \mathbb{R}^d$ and $f : D \to \mathbb{R}$ we denote the *partial derivative* w.r.t. the n–th variable as ∂_{x_n}. We define the *divergence* and the *gradient* as usual

$$\operatorname{div}(f) := \sum_{i=1}^{d} \partial_{x_i} f_i,$$
$$\nabla f := (\partial_{x_1} f, \dots, \partial_{x_n} f)^T.$$

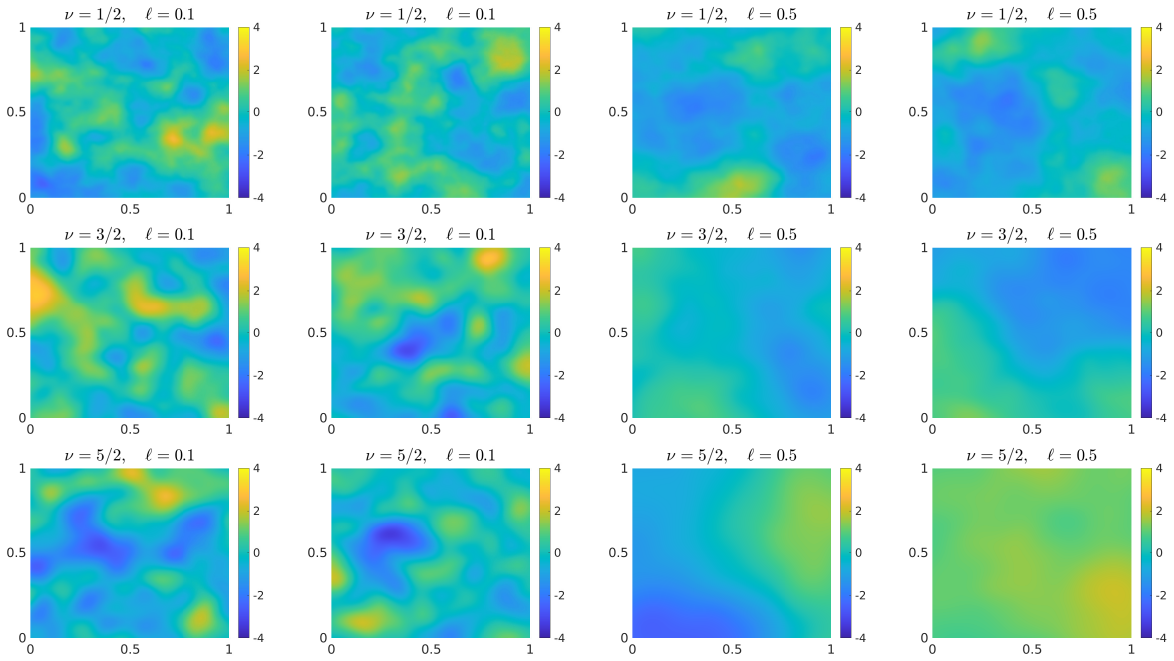

Figure 2.2: Samples of a Gaussian mean zero random field with Whittle–Matérn covariance in $D = [0,1]^2$ for different values of the smoothness $\nu \in \{1/2, 3/2, 5/2\}$ and correlation lengths $\ell \in \{0.1, 0.5\}$. Each image is an independent realization of the random field. We generated the plots by truncating the KLE after $N = 100$ terms.

We model heat transfer through a material with the help of Poisson's equation. The mathematical description for the temperature y is the solution of a *partial differential equation* (PDE). The physical interpretation is that a constant heat source, say a chemical reaction or the heating of a metal rod, has enough time to conduct through a material [7, Chapter 7.1]. The steady state temperature is then described with Poisson's equation, where we first introduce the strong formulation.

Example 2.30 (Strong Poisson equation)**.** Let $D \subseteq \mathbb{R}^d$ be a bounded domain. Then the *strong Poisson equation* is

$$\begin{aligned} -\operatorname{div}(a(x)\nabla y(x)) &= u(x), \qquad x \in D, \\ y(x) &= g(x), \qquad x \in \partial D. \end{aligned} \tag{2.8}$$

Let us describe the quantities and their physical meaning:

- The domain D describes the volume of a material.
- The function $y : D \to \mathbb{R}$ is the temperature in the material.
- The forcing function u controls how much heat is generated or lost inside the domain.
- The diffusion $a > 0$ determines how fast the heat travels through the material.
- The temperature at the boundary has fixed value g. ◇

This model problem is also used in groundwater flow, where it models subsurface flow. Here Darcy's law is combined with the continuity equation to obtain Poisson's equation, see [30, 40, 141]. The existence and uniqueness of solutions of strong PDE formulations is

often difficult to prove. Instead, we look at weak formulations. We further randomize the diffusion a which leads to a pathwise formulation. Weak formulations are often obtained by multiplying the PDE with suitably smooth test functions and using integration by parts. We denote with $H_0^1(D) \subseteq L^2(D)$ the *Sobolev space* consisting of functions with weak first order derivative in $L^2(D)$ and zero trace [50, Section 5]. We equip this space with the norm $\|y\|_{H_0^1(D)} := \|\nabla y\|_{L^2(D)}$.

Definition 2.31 (Pathwise weak elliptic PDE)**.** Let $D \subseteq \mathbb{R}^d$ be a bounded domain, $u \in L^2(D)$ and $g := 0$. We call $y \in H_0^1(D)$ a *weak solution* of (2.8) if for all functions $v \in H_0^1(D)$

$$(a\nabla y, \nabla v)_{L^2(D)} = (u, v)_{L^2(D)}.$$

We call y a *pathwise weak solution* if for $\mathbb{P}$–almost all $\omega \in \Omega$ the function $y(\omega) \in H_0^1(D)$ and for all functions $v \in H_0^1(D)$

$$(a(\omega)\nabla y(\omega), \nabla v)_{L^2(D)} = (u, v)_{L^2(D)}. \tag{2.9}$$

◇

This model is well–studied and often used as a baseline for numerical experiments [25, 31, 63, 87, 134], which is why we also use it. It is of course possible to further randomize some appearing quantities. Equation (2.9) can be generalized to account for random boundary values $g \neq 0$ or a random right–hand side u. This is done in [26, 134]. We are often not directly interested in the solution y but rather some quantity that we derive from it. As an example, we might define the QoI as average over a subset of the whole domain $D_{\text{obs}} \subseteq D$

$$Z(\omega) := \frac{1}{|D_{\text{obs}}|} \int_{D_{\text{obs}}} y(\omega, x) dx.$$

Properties of the solution. We have to make some assumptions on the diffusion coefficient a to ensure existence and uniqueness of the solution y. These assumptions are a reformulation of the assumptions in [26] and [134]. The authors of [26] assume a domain $D \in C^2$, i.e. with smooth boundary, and [134] extend the results to piecewise polygonal domains.

Assumption 2.32 (Properties of the diffusion a)**.** The diffusion coefficient a satisfies the following three properties:

- There exists a $t \in (0, 1]$ such that $\mathbb{P}$–almost surely realizations of a are in $C^t(\overline{D})$.
- The diffusion a satisfies the *pathwise ellipticity bound* such that for almost all ω

  $$0 < a_{\min}(\omega) \leq a(x, \omega) \leq a_{\max}(\omega) < +\infty \quad \text{for almost all } x \in \overline{D}, \tag{2.10}$$

  where $a_{\min}$ and $a_{\max}$ are random variables.
- For all $s_1, s_2 \in \mathbb{R}$ the bounds on the diffusion coefficient satisfy $a_{\min}^{s_1}, a_{\max}^{s_2} \in L^2$. ◇

Assumption 2.32 ensures that $a \in L^p(\Omega, C^t(\overline{D}))$ for all $p \in [1, +\infty)$. The lognormal diffusion coefficient satisfies this assumption.

Lemma 2.33 (Lognormal diffusion coefficient)**.** Let $\kappa \sim N(0, \mathcal{C})$ with covariance such that the kernel $k : \mathbb{R}_{\geq 0} \to \mathbb{R}$ defined as

$$k(\|x - y\|) := \mathbb{C}\mathrm{ov}[\kappa(x), \kappa(y)] \quad \text{for all } x, y \in D$$

is Lipschitz continuous. Then the lognormal diffusion coefficient

$$a := \exp(\kappa)$$

satisfies Assumption 2.32 for all t with $t < 1/2$. The Whittle–Matérn covariance kernel satisfies this assumption for the smoothness $\nu = 1/2$ for all $t < 1/2$ and for smoothness $\nu = 3/2$ and $\nu = 5/2$ with $t = 1$.

Proof. The result can be deduced from [25, Section 2] and by showing that the Whittle–Matérn covariance kernel is Lipschitz continuous. Therefore, we only outline the main idea. First, Kolmogorov's Theorem [34, Theorem 3.5] is used to verify that there exists a version of a whose realizations are Hölder continuous and thus continuous. We are then able to define the bounds

$$a_{\min}(\omega) := \min_{x \in \overline{D}} \exp(\kappa(x, \omega)), \qquad a_{\max}(\omega) := \max_{x \in \overline{D}} \exp(\kappa(x, \omega))$$

and (2.10) is satisfied since the exponential maps to the positive reals. The Fernique Theorem [34, Section 2.2] can now be used to bound the moments of $a_{\min}$, $a_{\max}$ and their inverse. This can be done similar to [25, Proposition 2.3] and [26, Proposition 2.4]. The smoothness of the sample paths for $\nu = 3/2$ and $\nu = 5/2$ follows from [111, Corollary 4.4]. The assumptions of this corollary are satisfied, since the covariance kernel is twice continuously differentiable with Hölder continuous derivative. This can be deduced from Definition 2.26 or from the expansion [128, Chapter 2, Equation (15)]. □

We use standard PDE theory [50, Chapter 6] to show the existence and uniqueness of a pathwise weak solution as well as pathwise bounds under some mild assumptions.

Theorem 2.34 (Existence, uniqueness and regularity of pathwise weak solutions)**.** Let $D \subseteq \mathbb{R}^d$ be a bounded Lipschitz domain and let Assumption 2.32 be true for $t \in (0, 1]$. Then there exists a unique pathwise weak solution y of (2.9) such that

$$\|y(\omega)\|_{H_0^1(D)} \leq c_1(\omega)\|u\|_{L^2(D)}, \tag{2.11}$$

$$\|y(\omega)\|_{L^2(D)} \leq c_2(\omega)\|u\|_{L^2(D)}. \tag{2.12}$$

Furthermore, for all $0 < s < t$ except $s = 1/2$ we have $y(\omega) \in H^{1+s}(D)$ and the bound

$$\|y(\omega)\|_{H^{1+s}(D)} \leq c_3(\omega)\|u\|_{L^2(D)}. \tag{2.13}$$

The random variables c_1, c_2, $c_3 \in L^p$ for every $p \in [1, +\infty)$. For $t = 1$ the statement holds with $s = 1$.

Proof. The existence, uniqueness and (2.11) is found in [25, Proposition 2.4] and is a result of the classical Lax–Milgram Lemma [50, Section 6.2.1]. The use of Poincare's inequality then shows (2.12). The bound (2.13) is given in [26, Proposition 3.1]. The moments of c_1, c_2, c_3 are bounded according to [26, Theorem 3.4]. □

A further computation shows that y is actually a well–defined random variable.

Lemma 2.35 (y is a random variable)**.** Let the assumptions of Theorem 2.34 be true. Then the unique pathwise weak solution y of (2.9) is a random variable $y \in L^p(\Omega, H_0^1(D) \cap H^{1+s}(D))$ for all $p \in [1, +\infty)$.

Proof. We verify that y is measurable by showing that y is locally Lipschitz continuous w.r.t. a since continuous functions are measurable, see Lemma 2.4. We view the solution y as function of the diffusion coefficient a

$$y : \{a \in L^\infty(\overline{D}) \,|\, \text{There exists } a_{\min} > 0 : \, a(x) \geq a_{\min} > 0 \text{ for a.a. } x \in \overline{D}\} \to H_0^1(D),$$

where a.a. is the abbreviation for almost all. Let $a, \widetilde{a}$ be two diffusion coefficients and $y := y(a)$, $\widetilde{y} := y(\widetilde{a})$ the respective solutions. A computation now shows

$$\|y - \widetilde{y}\|_{H_0^1(D)}^2 = \|\nabla y - \nabla \widetilde{y}\|_{L^2(D)}^2 \leq \frac{1}{a_{\min}} (a(\nabla y - \nabla \widetilde{y}), \nabla y - \nabla \widetilde{y})_{L^2(D)}.$$

We split this expression and use the weak formulation (2.9) with $v = y - \widetilde{y}$ once for y and once for $\widetilde{y}$ to conclude

$$\begin{aligned}
(a(\nabla y - \nabla \widetilde{y}), \nabla y - \nabla \widetilde{y})_{L^2(D)} &= (a \nabla y, \nabla y - \nabla \widetilde{y})_{L^2(D)} - (a \nabla \widetilde{y}, \nabla y - \nabla \widetilde{y})_{L^2(D)} \\
&= (u, y - \widetilde{y})_{L^2(D)} - (a \nabla \widetilde{y}, \nabla y - \nabla \widetilde{y})_{L^2(D)} \\
&= (\widetilde{a} \nabla \widetilde{y}, \nabla y - \nabla \widetilde{y})_{L^2(D)} - (a \nabla \widetilde{y}, \nabla y - \nabla \widetilde{y})_{L^2(D)} \\
&\leq \|\widetilde{a} - a\|_{L^\infty} \|\nabla \widetilde{y}\|_{L^2(D)} \|\nabla y - \nabla \widetilde{y}\|_{L^2(D)}.
\end{aligned}$$

We now use Poincare's inequality to show the result

$$\|\nabla \widetilde{y}\|_{L^2(D)}^2 \leq \frac{1}{\widetilde{a}_{\min}} (\widetilde{a} \nabla \widetilde{y}, \nabla \widetilde{y})_{L^2(D)} = \frac{1}{\widetilde{a}_{\min}} (u, \widetilde{y})_{L^2(D)} \leq \frac{c}{\widetilde{a}_{\min}} \|u\|_{L^2(D)} \|\nabla \widetilde{y}\|_{L^2(D)}. \qquad \square$$

Finite element method. The numerical computation of y requires us to discretize the Sobolev space $H_0^1(D)$ to obtain a discrete formulation of (2.9). In this thesis we restrict ourselves to linear finite elements and polygonal domains D. We now basically follow [29], a further introduction for finite element spaces is given by Brenner [21]. For the pathwise formulation we use results from [25, 26, 134].

Definition 2.36 (Finite element mesh)**.** Let $D \subseteq \mathbb{R}^2$ be a bounded and polygonal domain. Then $\mathcal{T} := \{\tau_1, \ldots, \tau_N\}$ is an *admissible mesh* if

- The $\tau_n \subseteq \mathbb{R}^2$ are open triangles,
- D is the union of these triangles $\overline{D} = \bigcup_{n=1}^N \overline{\tau}_n$,
- The triangles are disjoint $\tau_n \cap \tau_j = \emptyset$ for $n \neq j$,
- Any face (vertex or edge) of $\overline{\tau}_n$ is a face of another triangle $\overline{\tau}_j$ or is a subset of the boundary ∂D.

The *mesh size* h is the diameter of the largest triangle $h := \max_{n \in \{1,\ldots,N\}} \text{diameter}(\tau_n)$. A sequence of triangulations $(\mathcal{T}_\ell)_{\ell=1}^\infty$ is called *shape-regular* if there exists a constant $c > 0$ such that for all $\ell \in \mathbb{N}$ and all $\tau_n \in \mathcal{T}_\ell$

$$\frac{\text{inscribedradius}(\tau_n)}{\text{diameter}(\tau_n)} \geq c,$$

where inscribedradius is the radius of the largest inscribed circle of τ_n. $\diamond$

Definition 2.37 (Linear finite elements)**.** Let $\mathcal{T}_\ell$ be an admissible mesh. Then the space of *linear finite elements* is

$$V_\ell^{\mathrm{FE}} := \big\{ v \in C(\overline{D}) \,|\, v|_{\tau_n} \text{ is affine linear for all } \tau_n \in \mathcal{T}_\ell \text{ and } v|_{\partial D} = 0 \big\}.$$

The space of linear finite elements is *conforming* $V_\ell^{\mathrm{FE}} \subseteq H_0^1(D)$ for all $\ell \in \mathbb{N}$, see Brenner [21, Chapter 3]. ◇

The finite element method replaces $H_0^1(D)$ with V_ℓ^{FE} in the weak formulation (2.9).

Definition 2.38 (Pathwise discrete solution)**.** We call y_ℓ a *pathwise discrete solution* if for $\mathbb{P}$-almost all $\omega \in \Omega$ the function $y_\ell(\omega) \in V_\ell^{\mathrm{FE}}$ and for all $v_\ell \in V_\ell^{\mathrm{FE}}$

$$(a(\omega)\nabla y_\ell(\omega), \nabla v_\ell)_{L^2(D)} = (u, v_\ell)_{L^2(D)}. \tag{2.14}$$

◇

We write down the discrete analogon for the existence, uniqueness and boundedness of pathwise weak solutions.

Theorem 2.39 (Existence and uniqueness of pathwise discrete solutions)**.** Let Assumption 2.32 be true. Then there exists a unique pathwise discrete solution y_ℓ of (2.14) and it is bounded by

$$\begin{aligned} \|y_\ell(\omega)\|_{H_0^1(D)} &\le c_1(\omega)\|u\|_{L^2(D)}, \\ \|y_\ell(\omega)\|_{L^2(D)} &\le c_2(\omega)\|u\|_{L^2(D)}. \end{aligned}$$

Furthermore, $c_1, c_2 \in L^p$ for all $p \in [1, +\infty)$ and $y_\ell \in L^p(\Omega, V_\ell^{\mathrm{FE}})$.

Proof. The proof is analogous to the proof of Theorem 2.34 and Lemma 2.35. □

The approximation quality of the solution is summarized in the next theorem. The convergence depends crucially on the mesh size h and the convergence rate on the smoothness of the diffusion coefficient.

Theorem 2.40 (Finite element error estimate)**.** Let Assumption 2.32 be true for $t \in (0, 1]$. Then for all $0 < s < t$ except $s = 1/2$ the approximation error is bounded

$$\begin{aligned} \|y_\ell(\omega) - y(\omega)\|_{H_0^1(D)} &\le c_1(\omega) h_\ell^s \|u\|_{L^2(D)}, \\ \|y_\ell(\omega) - y(\omega)\|_{L^2(D)} &\le c_2(\omega) h_\ell^{2s} \|u\|_{L^2(D)}. \end{aligned}$$

Furthermore, the constants $c_1, c_2 \in L^p$ for all $p \in [1, +\infty)$ and thus

$$\begin{aligned} \mathbb{E}\Big[\|y_\ell - y\|_{H_0^1(D)}^p \Big] &\le c h_\ell^{ps} \|u\|_{L^2(D)}^p, \\ \mathbb{E}\Big[\|y_\ell - y\|_{L^2(D)}^p \Big] &\le c h_\ell^{2ps} \|u\|_{L^2(D)}^p. \end{aligned} \tag{2.15}$$

For $t = 1$ the statement of this theorem holds with $s = 1$.

Proof. The proof for the estimates in the $H_0^1(D)$ can be found in [26, Theorem 3.9] and the Aubin–Nitsche trick is used to obtain estimates for $L^2(D)$, see [26, Corollary 3.10]. □

We give the reader examples in Figure 2.3 of the solution y for different diffusion coefficients in $d = 2$. The general behaviour of the solution y is that if the diffusion is small in a certain area, then the temperature generated by u accumulates in this area. For large diffusion values the heat travels very fast to the boundary where we fixed the temperature to zero and thus the temperature in this area is comparatively low. This also explains why the temperature in the center is often largest, since the heat has to travel the farthest distance to the boundary.

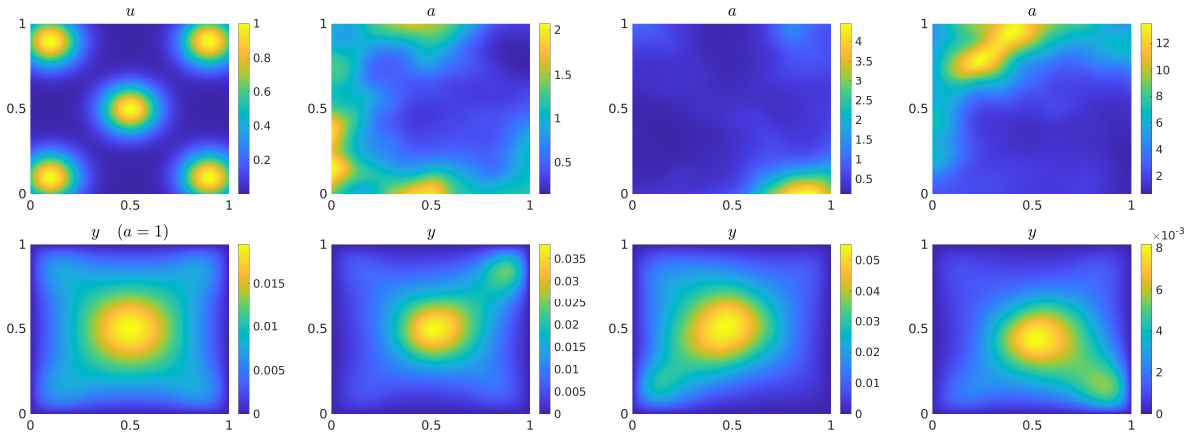

Figure 2.3: The left top image shows the function u which models heat generation at five points in some material. The bottom left image is the temperature y with a diffusion coefficient $a = 1$. The other images show the random field on top, which are realizations of $a := \exp(\kappa)$, where $\kappa \sim N(0, \mathcal{C})$ and $\mathcal{C}$ is obtained from the Whittle–Matérn kernel with $\nu := 3/2$ and $\ell := 1/2$. The three bottom images are the solutions y obtained from the diffusion a above.

The convergence rates of y naturally extends to a QoI Z that is defined as an average value over a subdomain $D_{\text{obs}} \subseteq D$

$$\begin{aligned} Z(\omega) &:= \frac{1}{|D_{\text{obs}}|} \int_{D_{\text{obs}}} y(\omega, x) dx, \\ Z_\ell(\omega) &:= \frac{1}{|D_{\text{obs}}|} \int_{D_{\text{obs}}} y_\ell(\omega, x) dx \quad \text{for all } \ell \in \mathbb{N}. \end{aligned} \tag{2.16}$$

We write down the corresponding error estimate.

Corollary 2.41 (Finite element error estimate for Z)**.** Let Assumption 2.32 be with $t = 1$. Then $Z, Z_1, Z_2, \cdots \in L^p$ for all $p \in [1, +\infty)$ and the approximation error is bounded

$$\mathbb{E}\big[\|Z_\ell - Z\|^2\big] \leq c h_\ell^4. \tag{2.17}$$

In particular, we have the following errors for the mean and variance for all $\ell \in \mathbb{N}$

$$\|\mathbb{E}[Z_\ell] - \mathbb{E}[Z]\| \leq c h_\ell^2, \tag{2.18}$$

$$\mathbb{V}[Z_\ell - Z] \leq c h_\ell^4. \tag{2.19}$$

Proof. The result (2.17) follows from Jensen's inequality and (2.15) in Theorem 2.40

$$\begin{aligned} \mathbb{E}\big[\|Z_\ell - Z\|^2\big] &\leq \frac{1}{|D_{\text{obs}}|^2} \mathbb{E}\left[\left(\int_{D_{\text{obs}}} |y_\ell(x) - y(x)| dx\right)^2\right] \leq \frac{1}{|D_{\text{obs}}|^2} \mathbb{E}\left[\int_D |y_\ell(x) - y(x)|^2 dx\right] \\ &\leq \frac{1}{|D_{\text{obs}}|^2} \mathbb{E}\Big[\|y_\ell - y\|^2_{L^2(D)}\Big] \leq c h_\ell^4. \end{aligned}$$

We show (2.18) by pulling out the expectation, using Jensen's inequality and (2.17)

$$\|\mathbb{E}[Z_\ell] - \mathbb{E}[Z]\|^2 \leq \mathbb{E}\big[\|Z_\ell - Z\|^2\big] \leq c h_\ell^4.$$

We use the fact that $\mathbb{E}[Z_\ell - Z]$ is the deterministic constant that minimizes the quadratic deviation from $Z_\ell - Z$ and use (2.17) to show (2.19)

$$\begin{aligned}\mathbb{V}[Z_\ell - Z] &= \mathbb{E}\big[\|Z_\ell - Z - \mathbb{E}[Z_\ell - Z]\|^2\big] = \min_{\mu\in\mathbb{R}} \mathbb{E}\big[\|Z_\ell - Z - \mu\|^2\big] \\ &\leq \mathbb{E}\big[\|Z_\ell - Z - 0\|^2\big] \leq c h_\ell^4. \end{aligned}$$ □

Numerical implementation. We outline how to numerically compute a solution $y_\ell(\omega) \in V_\ell^{\mathrm{FE}}$ of (2.14). Let $(\varphi_n)_{n=1}^N$ be any finite dimensional basis of V_ℓ^{FE} and express the solution of $y_\ell(\omega)$ as linear combination of the basis functions of this space

$$y_\ell(\omega) = \sum_{n=1}^N \beta_n(\omega)\varphi_n.$$

Then test (2.14) with all test functions $\varphi_n \in V_\ell^{\mathrm{FE}}$ to obtain the pathwise linear system of equations of the form such that for $\mathbb{P}$–almost all $\omega \in \Omega$

$$A(\omega)\beta(\omega) = b. \tag{2.20}$$

The *stiffness matrix* $A(\omega)$ is then

$$A(\omega) := (A_{nj}(\omega))_{n,j=1}^N := ((a(\omega)\nabla\varphi_n, \nabla\varphi_j)_{L^2(D)})_{n,j=1}^N \in \mathbb{R}^{N\times N}. \tag{2.21}$$

The right–hand side or *load vector* b is given as follows

$$b := (b_n)_{n=1}^N := ((u, \varphi_n)_{L^2(D)})_{n=1}^N \in \mathbb{R}^N. \tag{2.22}$$

The solution *solution vector* $\beta(\omega) \in \mathbb{R}^N$ is the vector of coefficients for $y_\ell(\omega)$ satisfying (2.20)

$$\beta(\omega) := (\beta_n(\omega))_{n=1}^N \in \mathbb{R}^N.$$

It is often required to compute the $L^2(D)$-norm of y_ℓ which can be achieved with the help of the *mass matrix*

$$M := (M_{nj})_{n,j=1}^N := ((\varphi_n, \varphi_j))_{n,j=1}^N$$

such that $\|y_\ell\|_{L^2(D)}^2 = \beta^T M \beta$.

Remark 2.42 (Quadrature rules)**.** It is often necessary to use a quadrature rule to compute the stiffness matrix A in (2.21) and the load vector b in (2.22). This introduces additional errors that may worsen the error rates of Theorem 2.40 if $a(\omega)$ is not sufficiently smooth, see [26, Section 3.3]. ◇

Remark 2.43 (Errors in the diffusion coefficient)**.** We also replace the diffusion coefficient a by a truncated KLE in our numerical experiments. We do not state the error that gets introduced from using a truncation and instead refer to Charrier [25, 26]. ◇

We summarize the method to compute samples of y_ℓ. First, we have to compute the KLE of the diffusion coefficient such that we are able to cheaply generate samples of a. Afterwards, we obtain a single sample by the following steps:

1. Compute a realization of the diffusion coefficient $a(\omega)$.
2. Compute the stiffness matrix $A(\omega)$ according to (2.21) using some quadrature rule.
3. Compute the load vector b according to (2.22) using some quadrature rule.
4. Solve the system (2.20) to obtain the coefficients $\beta(\omega)$ of $y_\ell(\omega) \in V_\ell^{\mathrm{FE}}$.

The most expensive step is often to solve the system (2.20). In particular, if the basis functions $(\varphi_n)_{n=1}^N$ have full support over D, then $A(\omega)$ is a dense matrix and thus the costs to directly solve the system (2.20) is of order $\mathcal{O}(N^3)$. A standard method to reduce the complexity is to use the *nodal basis*, which are functions that have local support. Then the stiffness matrix A is sparse, which allows us to efficiently solve the system (2.20). Multigrid methods [137, Section 3] are able to compute a solution with accuracy ε with costs $\mathcal{O}(N \log(\varepsilon))$, which is optimal up to logarithmic factors.
We remark that we view two solutions on two different grids y_{ℓ_1} and y_{ℓ_2} as functions in $H_0^1(D)$. This ensures, for example, that the expression $y_{\ell_1} - y_{\ell_2}$ is well defined, however the corresponding expression for the vectors $\beta^{\ell_1} - \beta^{\ell_2}$ is in general not well defined. We overcome this obstacle if $V_{\ell_1}^{\mathrm{FE}} \subseteq V_{\ell_2}^{\mathrm{FE}}$ with a linear *prolongation operator* $P_{\ell_2,\ell_1}^{\mathrm{FE}}$ such that $P_{\ell_2,\ell_1}^{\mathrm{FE}} y_{\ell_1}$ are the coefficients of y_{ℓ_1} in the nodal basis of $V_{\ell_2}^{\mathrm{FE}}$. That is, the coefficients of the difference $y_{\ell_1} - y_{\ell_2}$ in the nodal basis of $V_{\ell_2}^{\mathrm{FE}}$ is computed as follows

$$P_{\ell_2,\ell_1}^{\mathrm{FE}} \beta^{\ell_1} - \beta^{\ell_2}.$$

Chapter 3

Estimation and variance reduction

The goal of this chapter is to give an overview for the approximation of the mean of a random variable $Z \in L^2$

$$\mathbb{E}[Z] = \int_\Omega Z(\omega) d\mathbb{P}(\omega). \tag{3.1}$$

We divide quadrature methods into deterministic and sampling based quadrature and list common methods for both.

- *Deterministic quadrature*: These methods often discretize $\mathbb{P}$ and approximate (3.1) as a weighted sum of evaluations of Z at quadrature points. Classical quadrature rules like the *trapezoidal rule* or *Gaussian quadrature* (see [35, Chapter 5] or [38]) work well in one dimension and reach a high accuracy. The straightforward extension to multiple dimensions is called *tensor product quadrature* and applies these rules for every dimension. This is prohibitively expensive in high dimensions, since the cost is exponential in the number of dimensions. A potential improvement are *sparse grids* [24] where we do not use a full tensor grid but only a carefully chosen subset of grid points. In the context of random PDEs this is often combined with *collocation methods* [102]. Sparse grids require mixed regularity of the integrand w.r.t. the dimensions, which may not always be satisfied. The *stochastic Galerkin method* [55] or [92, Chapter 9] discretizes the probability space with the help of polynomials. The approach is similar to the finite element discretization of $H_0^1(D)$ and may result in a high dimensional problem. The drawback of this method is that it requires the solution of a large linear system and is intrusive. Another well–known approach is *quasi-Monte Carlo* [46], [90, Section 5,6], which uses the same quadrature weights as Monte Carlo with deterministic evaluation points obtained from a low-discrepancy sequence. This also requires the reformulation of (3.1) as integral over $\mathbb{R}^s$.

- *Sampling based quadrature*: These methods combine samples of Z. We explain general results and introduce some common notation in Section 3.1. We examine the classical Monte Carlo estimator in Section 3.2 adapted for the PDE setting with random coefficients. Control Variates are a variance reduction technique and we study them in Section 3.3. The specific implementations of Control Variates are Multifidelity Monte Carlo in Section 3.4 and Approximate Control Variates in Section 3.5. Another popular variance reduction method in the context of hierarchical models is Multilevel Monte Carlo in Section 3.6. We finish this chapter with a brief review of other sampling based methods in Section 3.7.

The thesis is mainly concerned with sampling based quadrature and we do not treat deterministic quadrature any further. Most results in this chapter are well known and a repetition or reformulation adapted for the goal (3.1). Throughout the rest of thesis if not mentioned otherwise, we assume that Z is a real–valued random variable. Furthermore, we assume that $Z \in L^2$ such that Z has finite first and second moments. Generalizations for $Z \in L^1$ with $Z \notin L^2$ are possible but complicate the analysis. For example, the strong law of large numbers [79, Theorem 5.17] shows that the sample average converges

almost surely to the mean under the weak assumption $Z \in L^1$. However, for a rate of convergence w.r.t. the number of samples we require higher regularity and $Z \in L^2$ is sufficient to obtain a usable rate.
We call the next assumption *General Assumption*, since we assume it throughout the thesis without further mentioning it.

General Assumption 3.1. Let $Z, Z_1, Z_2, \dots, Z_L \in L^2$ be real–valued random variables and let Z_ℓ have positive cost $w_\ell > 0$ for all $\ell \in \{1, \dots, L\}$. ◇

Challenges. Let us briefly demonstrate the challenges that arise for the computation of the mean of a random variable.

Example 3.2 (Diffusion process)**.** Let us define the QoI as in (2.16)

$$Z(\omega) := \frac{1}{|D_{\text{obs}}|} \int_{D_{\text{obs}}} y(a(\omega), x) dx,$$

where $D_{\text{obs}} \subseteq D$ and y is a pathwise weak solution of the elliptic PDE in Definition 2.31 with diffusion coefficient

$$a(x, \omega) := \sum_{n=1}^{\infty} \sqrt{\lambda_n} \xi_n(\omega) \psi_n(x). \tag{3.2}$$

We denote the truncation of this expansion at N with a_N and obtain an approximation of the integral

$$\mathbb{E}[Z] = \int_\Omega Z(a(\omega)) d\mathbb{P}(\omega) \approx \int_{\mathbb{R}^N} Z(a_N(y)) p_N(y) dy,$$

where p_N is the probability density function of the random vector $(\xi_1, \dots, \xi_N)^T$. We are not able to compute the solution y or Z analytically and thus replace them with the approximation $Z_L \approx Z$

$$\mathbb{E}[Z] \approx \int_{\mathbb{R}^N} Z(a_N(y)) p_N(y) dy \approx \int_{\mathbb{R}^N} Z_L(a_N(y)) p_N(y) dy. \tag{3.3}$$

◇

We list the three main challenges to compute the mean and list reasons why we should be able to overcome them.

- **Curse of dimensionality:** The probability space is often high–dimensional, e.g. in (3.3) of dimension N, and thus requires high-dimensional quadrature. The curse of dimensionally describes the phenomena that the costs grow rapidly or exponentially w.r.t. the dimension N of the stochastic space. However, as Lemma 2.28 shows the error of dropping the random variable ξ_n from the KLE in (3.2) is small if n is large, which translates to a small error for Z. Although the original problem (3.1) requires the evaluation of an infinite dimensional integral, the effect of large dimensions n on Z decays. This intuition is formalized with the variance since $\mathbb{V}[Z] < +\infty$ is bounded independently of the truncation index N.

- **Discretization:** We have approximations $Z_1, Z_2, \dots$ of Z at varying degrees of accuracy. In Example 3.2 we use a finite element discretization for y and Theorem 2.40 shows that the error decreases depending on the mesh size h. As the

mesh size decreases the quadrature error decreases, however the required computational resources increase significantly. The coarse discretizations have a larger error than the fine discretizations but are much cheaper to compute. Since the coarse grids contain some information it is conceptionally advantageous to include them for estimating the mean.

- **Cost balancing:** The quadrature and the discretization error have to be balanced to achieve a computationally tractable solution. If the discretization is too fine we are only able to compute a few evaluations of Z_ℓ and the quadrature error of $\mathbb{E}[Z]$ is large. On the other hand if the discretization Z_ℓ is coarse and cheap the quadrature error is small but the discretization error is large. Furthermore, a method that uses discretizations $Z_1, Z_2, \dots$ with different costs and accuracies has to balance the computational effort between these models.

3.1 Sampling based estimation

The goal of this section is to define some common terminology and explain the reason for the sampling based estimation of the mean. Sampling based estimators define a random variable $\widehat{\mu}$ that approximates the mean

$$\widehat{\mu} \approx \mathbb{E}[Z].$$

We understand $\approx$ in the mean square error sense, which is the average squared deviation from the quantity $\mathbb{E}[Z]$. We further define the bias that denotes how far $\mathbb{E}[\widehat{\mu}]$ deviates from $\mathbb{E}[Z]$. This definition is rigorous if we assume $\widehat{\mu} \in L^2$, which is the case for all examined estimators in this thesis.

Definition 3.3 (Bias, mean square error)**.** We define the *bias* and *mean square error* (MSE) of an estimator $\widehat{\mu}$ as

$$\begin{aligned}\mathrm{Bias}(\widehat{\mu}) &:= \mathrm{Bias}(\widehat{\mu}, Z) := \|\mathbb{E}[\widehat{\mu}] - \mathbb{E}[Z]\|,\\ \mathrm{MSE}(\widehat{\mu}) &:= \mathrm{MSE}(\widehat{\mu}, Z) := \mathbb{E}\big[\|\widehat{\mu} - \mathbb{E}[Z]\|^2\big],\end{aligned}$$

where $\|\cdot\|$ is a suitable norm. We call an estimator $\widehat{\mu}$ *unbiased* if $\mathrm{Bias}(\widehat{\mu}) = 0$ and thus $\mathbb{E}[\widehat{\mu}] = \mathbb{E}[Z]$. ◇

The bias and MSE are always defined w.r.t. the expectation $\mathbb{E}[Z]$ that we want to estimate. We often abbreviate $\mathrm{Bias}(\widehat{\mu})$ and $\mathrm{MSE}(\widehat{\mu})$ without the explicit Z if there is no ambiguity. It is well known that the MSE satisfies a bias-variance decomposition and is equal to the squared bias plus the variance of the estimator. For unbiased estimators the MSE is equal to the variance.

Theorem 3.4 (Bias-variance decomposition)**.** The MSE admits the decomposition

$$\mathrm{MSE}(\widehat{\mu}) = \mathrm{Bias}(\widehat{\mu})^2 + \mathbb{V}[\widehat{\mu}].$$

In particular, for unbiased estimators $\widehat{\mu}$ the MSE is equal to the variance

$$\mathrm{MSE}(\widehat{\mu}) = \mathbb{V}[\widehat{\mu}].$$

Proof. We insert $\mathbb{E}[\widehat{\mu}]$ and use the properties of the inner product

$$\begin{aligned}\mathbb{E}\big[\|\widehat{\mu} - \mathbb{E}[Z]\|^2\big] &= \mathbb{E}[\|\widehat{\mu} - \mathbb{E}[\widehat{\mu}] + \mathbb{E}[\widehat{\mu}] - \mathbb{E}[Z]\|^2] \\ &= \mathbb{E}\big[\|\mathbb{E}[\widehat{\mu}] - \mathbb{E}[Z]\|^2\big] + \mathbb{E}\big[\|\widehat{\mu} - \mathbb{E}[\widehat{\mu}]\|^2\big] \\ &\quad + 2\mathbb{E}[(\widehat{\mu} - \mathbb{E}[\widehat{\mu}], \mathbb{E}[\widehat{\mu}] - \mathbb{E}[Z])].\end{aligned}$$

The first term is the squared bias and the second term is the variance of the estimator $\widehat{\mu}$. For the third term we use the linearity of the inner product in the first argument and pull in the expectation. We then conclude that this term is zero. □

The bias-variance decomposition leads to a bound on the probability that $\widehat{\mu}$ deviates from the mean $\mathbb{E}[Z]$. Both the bias and the variance is small if the MSE is small and an application of Chebychev's inequality in Theorem 2.21 then bounds this probability.

Theorem 3.5 (MSE bound for probability of deviation)**.** For all positive ε with $\varepsilon^2 \geq 4\,\mathrm{MSE}(\widehat{\mu})$ the probability that $\widehat{\mu}$ deviates from $\mathbb{E}[Z]$ is bounded by

$$\mathbb{P}(\|\widehat{\mu} - \mathbb{E}[Z]\| > \varepsilon) \leq \frac{4}{\varepsilon^2}\,\mathrm{MSE}(\widehat{\mu}). \tag{3.4}$$

Proof. We insert $\mathbb{E}[\widehat{\mu}]$ and use the triangle inequality

$$\|\widehat{\mu} - \mathbb{E}[Z]\| \leq \|\widehat{\mu} - \mathbb{E}[\widehat{\mu}]\| + \mathrm{Bias}(\widehat{\mu}).$$

The expression on the left has equal or smaller probability to exceed ε than the expression on the right

$$\mathbb{P}(\|\widehat{\mu} - \mathbb{E}[Z]\| > \varepsilon) \leq \mathbb{P}(\|\widehat{\mu} - \mathbb{E}[\widehat{\mu}]\| + \mathrm{Bias}(\widehat{\mu}) > \varepsilon). \tag{3.5}$$

We use Theorem 3.4 and $\mathrm{MSE}(\widehat{\mu}) \leq \varepsilon^2/4$ to bound the bias

$$\mathrm{Bias}(\widehat{\mu}) \leq (\mathrm{Bias}(\widehat{\mu})^2 + \mathbb{V}[\widehat{\mu}])^{1/2} = \mathrm{MSE}(\widehat{\mu})^{1/2} \leq \varepsilon/2.$$

We combine this with Chebychev's inequality in Theorem 2.21 and $\mathbb{V}[\widehat{\mu}] \leq \mathrm{MSE}(\widehat{\mu})$ from Theorem 3.4 to conclude (3.4)

$$\mathbb{P}(\|\widehat{\mu} - \mathbb{E}[\widehat{\mu}]\| + \mathrm{Bias}(\widehat{\mu}) > \varepsilon) \leq \frac{\mathbb{V}[\widehat{\mu}]}{(\varepsilon - \mathrm{Bias}(\widehat{\mu}))^2} \leq \frac{4}{\varepsilon^2}\,\mathrm{MSE}(\widehat{\mu}).$$

□

The previous theorem allows us to derive confidence intervals for the estimation of the mean. It further allows us to upper bound the required MSE such that we do not deviate more than ε from the mean $\mathbb{E}[Z]$ with some prescribed probability δ.

Example 3.6 (Confidence intervals)**.** We determine the deviation $\varepsilon > 0$ such that for a prescribed confidence $\delta \in (0, 1)$

$$\mathbb{P}(\widehat{\mu} \in [-\varepsilon + \mathbb{E}[Z], \mathbb{E}[Z] + \varepsilon]) \geq \delta.$$

This is equivalent to demand that

$$\mathbb{P}(\|\widehat{\mu} - \mathbb{E}[Z]\| > \varepsilon) \leq 1 - \delta. \tag{3.6}$$

The use of (3.4) now gives the smallest deviation $\varepsilon = 2(\mathrm{MSE}(\widehat{\mu})/(1-\delta))^{1/2}$. ◇

Example 3.7 (Target MSE)**.** Let the deviation $\varepsilon > 0$ and confidence level $\delta \in (0,1)$ be fixed. We want to determine $\mathrm{MSE}(\widehat{\mu})$ to ensure that

$$\mathbb{P}(\widehat{\mu} \in [-\varepsilon + \mathbb{E}[Z], \mathbb{E}[Z] + \varepsilon]) \geq \delta.$$

Similarly to Example 3.6 we obtain (3.6) and thus (3.4) shows

$$1 - \delta \geq \frac{4}{\varepsilon^2} \mathrm{MSE}(\widehat{\mu}).$$

Therefore the estimator $\widehat{\mu}$ has to satisfy $\mathrm{MSE}(\widehat{\mu}) \leq (1-\delta)\varepsilon^2/4$. ◇

It is of course possible to improve these estimates if more knowledge of the underlying distribution of $\widehat{\mu}$ is available, see [45, Chapter 8] or [42, Chapter 24, 25], [122, Chapter 10] for the Gaussian case. Improved results are often obtained from concentration inequalities, see [19].

3.2 Monte Carlo

Exact sampling. In this section we review the basic properties of the classical Monte Carlo (MC) estimator. The results are well known and available in most introductory books on probability theory [77, 79]. The MC estimator or the sample average plays a central role for laws of large numbers, the central limit theorem and estimation in general.

Definition 3.8 (MC estimator)**.** For $m \in \mathbb{N}$ the *MC estimator* is the sample average

$$\widehat{\mu}^{\mathrm{MC}} := \frac{1}{m} \sum_{i=1}^{m} Z^i,$$

where the Z^i are m i.i.d. samples of Z. ◇

The MC estimator is unbiased and has variance inversely proportional to m. We provide a proof since the used techniques are useful for other estimators.

Lemma 3.9 (Bias, variance of MC estimator)**.** The MC estimator is an unbiased estimator for $\mathbb{E}[Z]$ with variance

$$\mathbb{V}\big[\widehat{\mu}^{\mathrm{MC}}\big] = \frac{\mathbb{V}[Z]}{m}. \tag{3.7}$$

Proof. The unbiasedness is a direct consequence of the linearity of the expectation and that all samples have the same distribution as Z. The expression for the variance crucially requires that $Z^1, \dots, Z^m$ are pairwise uncorrelated

$$\mathbb{V}\big[\widehat{\mu}^{\mathrm{MC}}\big] = \frac{1}{m^2} \sum_{i,j=1}^{m} \mathbb{C}\mathrm{ov}\big[Z^i, Z^j\big] = \frac{1}{m^2} \sum_{i,j=1}^{m} \begin{cases} \mathbb{V}[Z^i], & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

We use $\mathbb{V}[Z^i] = \mathbb{V}[Z]$ for $i \in \{1, \dots, m\}$ to conclude (3.7). □

The convergence rate for the MC estimator is thus of order $\mathcal{O}(m^{-1/2})$ for the *root mean square error* (RMSE)

$$\mathrm{RMSE}(\widehat{\mu}^{\mathrm{MC}}) := \mathrm{MSE}(\widehat{\mu}^{\mathrm{MC}})^{1/2} = \frac{\mathbb{V}[Z]^{1/2}}{m^{1/2}}.$$

The rate $\mathcal{O}(m^{-1/2})$ is rather slow compared to deterministic quadrature. However, the convergence rate is independent of the dimension of the underlying probability space, does not suffer from the curse of dimensionality and does not require any regularity assumptions except for $Z \in L^2$.
The cost to evaluate the MC estimator is the cost of a single samples times the number of samples. A consequence of Lemma 3.9 is that to half the RMSE we have to increase the number of samples by a factor of four. This also increases the cost by a factor of four.

Corollary 3.10 (Exact sampling cost)**.** The cost to achieve $\mathrm{MSE}(\widehat{\mu}^{\mathrm{MC}}) \leq \varepsilon^2$ is

$$\mathbb{W}\big[\widehat{\mu}^{\mathrm{MC}}\big] = m\mathbb{W}[Z] = \frac{1}{\varepsilon^2}\mathbb{V}[Z]\mathbb{W}[Z] = \mathcal{O}(\varepsilon^{-2}). \tag{3.8}$$

◇

We did not include the practically necessary ceiling of m in (3.8) since

$$m = \frac{1}{\varepsilon^2}\mathbb{V}[Z]$$

is typically not an integer. We continue to do so throughout the rest of this thesis and only mention the additional costs for ceiling if the obtained results are different. We remark that sampling based methods often require a lot of evaluations and thus ceiling the number of samples does not significantly increase the overall costs.

Discretized sampling. Corollary 3.10 makes the strong assumption that we are able to generate samples from the true QoI Z, which is often not possible. Therefore let $Z_1, Z_2, \ldots$ be a sequence of model discretizations such that the mean converges $\mathbb{E}[Z_\ell] \to \mathbb{E}[Z]$ for $\ell \to +\infty$. The bias-variance decomposition of the MSE now guides us how to proceed. First, select an appropriate fine level L such that the bias is smaller than a prescribed threshold. Afterwards apply the standard MC estimator for Z_L with enough samples to sufficiently decrease the variance. We denote this estimator with $\widehat{\mu}_L^{\mathrm{MC}}$ and provide a well–known result which is proved similarly to [56, Theorem 3.1].

Theorem 3.11 (Asymptotic cost)**.** Assume that for all $\ell \in \mathbb{N}$ the bias, variance and costs of the models Z_ℓ satisfy

$$\mathrm{Bias}(Z_\ell) \leq c2^{-\gamma_{\mathrm{Bias}}\ell}, \tag{3.9}$$
$$\mathbb{V}[Z_\ell] \leq c, \tag{3.10}$$
$$w_\ell \leq c2^{\gamma_{\mathrm{Cost}}\ell}. \tag{3.11}$$

Then for all $\varepsilon \in (0, 1/e]$ there exists a fine level $L \in \mathbb{N}$ and a number of samples $m \in \mathbb{N}$ such that the cost to achieve $\mathrm{MSE}(\widehat{\mu}_L^{\mathrm{MC}}) \leq \varepsilon^2$ is bounded

$$\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{MC}}\big] \leq c\varepsilon^{-2-\gamma_{\mathrm{Cost}}/\gamma_{\mathrm{Bias}}}. \tag{3.12}$$

Proof. We follow [56, Theorem 3.1]. Let $\varepsilon^2 > 0$ be fixed and set

$$L := -\log_2(\varepsilon)/\gamma_{\text{Bias}} + L_0$$

for $L_0 := L_0(\varepsilon^2) \geq 0$ such that $L \in \mathbb{N}$. We use (3.9) and assume L_0 is large enough to satisfy

$$\text{Bias}(\widehat{\mu}_L^{\text{MC}})^2 = \text{Bias}(Z_L)^2 \leq c2^{-2\gamma_{\text{Bias}}L} = c2^{-2\gamma_{\text{Bias}}L_0}\varepsilon^2 \leq \varepsilon^2/2.$$

The last inequality is achieved for L_0 bounded independently of ε^2. We define the number of samples as follows

$$m := \left\lceil \frac{2\sigma_L^2}{\varepsilon^2} \right\rceil \geq \frac{2\sigma_L^2}{\varepsilon^2}.$$

We combine this with the bias-variance decomposition and the expression for the variance (3.7) to obtain the bound for the MSE

$$\text{MSE}(\widehat{\mu}_L^{\text{MC}}) = \text{Bias}(\widehat{\mu}_L^{\text{MC}})^2 + \mathbb{V}\big[\widehat{\mu}_L^{\text{MC}}\big] \leq \frac{\varepsilon^2}{2} + \frac{\sigma_L^2}{m} \leq \varepsilon^2.$$

We apply the variance bound (3.10) and sample cost bound (3.11) to get the bound (3.12)

$$\begin{aligned}\mathbb{W}\big[\widehat{\mu}_L^{\text{MC}}\big] = mw_L = \left\lceil \frac{2\sigma_L^2}{\varepsilon^2} \right\rceil w_L &\leq c(\varepsilon^{-2}+1)2^{\gamma_{\text{Cost}}L} \leq c2^{\gamma_{\text{Cost}}L_0}\varepsilon^{-2-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} \\ &\leq c\varepsilon^{-2-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}.\end{aligned}$$

□

The cost $\mathcal{O}(\varepsilon^{-2-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}})$ is much larger than the cost $\mathcal{O}(\varepsilon^{-2})$ for exactly sampling from Z. This result is rather intuitive, since an increase of the accuracy not only requires more samples but also increases the cost per sample. Let us put this rate into perspective for the elliptic PDE with different physical dimensions.

Example 3.12 (PDE example)**.** Let Z be a QoI from an elliptic PDE similarly to (2.16). We now assume an optimistic costs increase equal to the degrees of freedom in a finite element mesh

$$w_\ell = 2^{d\ell},$$

where $\gamma_{\text{Cost}} = d$ is the dimension of the physical domain D. We use Corollary 2.41 to obtain the bias for level ℓ

$$\text{Bias}(Z_\ell) \leq ch_\ell^2.$$

This translates to $\gamma_{\text{Bias}} = 2$ for uniform mesh refinement with $h_\ell = c2^{-\ell}$. The cost is thus of order $\mathcal{O}(\varepsilon^{-2-d/2})$ and we give a summary in Table 3.1. The exact sampling cost of $\mathcal{O}(\varepsilon^{-2})$ leads to a cost increase of a factor of four if we half the RMSE. ◇

Dimension d	1	2	3	4	5	6
Cost	$\varepsilon^{-2.5}$	ε^{-3}	$\varepsilon^{-3.5}$	ε^{-4}	$\varepsilon^{-4.5}$	ε^{-5}
Cost increase	≈ 5.66	8	≈ 11.31	16	≈ 22.63	32

Table 3.1: Cost of the MC estimator w.r.t. to the dimension d of the domain D. The row "Cost increase" denotes the factor by which the total cost increases if we require half the RMSE ε. The QoI is (2.16) obtained from the elliptic PDE example.

Variance reduction. We outline the main idea for many variance reduction techniques. The cost of the MC estimator (3.8) is proportional to the variance $\mathbb{V}[Z]$ and we try to decrease this term without increasing the overall costs to much. As an example, let Y be a random variable with mean zero such that the MC estimator of $Z - Y$ is an unbiased estimator for $\mathbb{E}[Z]$. The costs to achieve a MSE of ε^2 are

$$\mathbb{W}\left[\frac{1}{m}\sum_{i=1}^{m}(Z^i - Y^i)\right] = \frac{\mathbb{V}[Z-Y]}{\varepsilon^2}\mathbb{W}[Z,Y],$$

where $\mathbb{W}[Z,Y]$ is the cost to compute a sample of (Z,Y). The cost compared to the standard MC estimator is reduced significantly if

$$\mathbb{V}[Z-Y]\mathbb{W}[Z,Y] \ll \mathbb{V}[Z]\mathbb{W}[Z].$$

This happens if $Z \approx Y$ and the cost to compute a sample of Y is much cheaper than Z. This approach is sensible since Y carries a lot of information required to compute $\mathbb{E}[Z]$ and we already know $\mathbb{E}[Y] = 0$. We give a mathematically rigorous interpretation later in Chapter 4. In any case, the estimators in the following sections all use some variation of this basic idea to construct a cheaper estimator.

3.3 Control Variates

Single control variates. The method of control variates is a well–known variance reduction technique and results of this section are available in [59, 60, 88, 123]. We start with the single control variate (CV) estimator and introduce multiple control variates later in this section. To construct the CV estimators we assume that the expectations $\mu_1, \dots, \mu_{L-1}$ are known.

Definition 3.13 (Single CV estimator)**.** For $m \in \mathbb{N}$ and $\beta \in \mathbb{R}$ the *CV estimator* is defined as

$$\widehat{\mu}_L^{\mathrm{CV}} := \frac{1}{m}\sum_{i=1}^{m} Z_L^i - \beta\left(\frac{1}{m}\sum_{i=1}^{m} Z_{L-1}^i - \mu_{L-1}\right).$$

We call the random variable Z_{L-1} *control variate*. ◇

The β does not influence the bias of the estimator, only its variance. Since we want to obtain an estimator that minimizes the MSE it makes sense to choose a β such that the variance is as small as possible. It turns out that the optimal value for β is unique under mild assumptions and independent of the number of samples.

Lemma 3.14 (Bias, variance)**.** The CV estimator is an unbiased estimator for μ_L with variance

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \frac{\sigma_L^2 - 2\beta C_{L,L-1} + \beta^2\sigma_{L-1}^2}{m}. \tag{3.13}$$

For $\sigma_{L-1}^2 > 0$ the variance minimizing coefficient is unique

$$\beta^* = \frac{C_{L,L-1}}{\sigma_{L-1}^2} \tag{3.14}$$

and the variance for this minimizer is

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = (1 - \rho_{L-1,L}^2)\frac{\sigma_L^2}{m}. \tag{3.15}$$

Proof. The unbiasedness of $\widehat{\mu}_L^{\mathrm{CV}}$ is obvious. The result (3.13) follows from the variance of sums of independent random variables in Lemma 2.18. The variance (3.13) as function of β is a parabola with positive leading coefficient since $\sigma_{L-1}^2 > 0$. Elementary calculus now shows (3.14) and (3.15). □

The variance of the CV estimator is never larger than the variance of the MC estimator, since $\rho_{L,L-1} \in [-1, 1]$ in (3.15), however, we have the additionally cost to compute a sample of Z_{L-1}. The next corollary states that the CV estimator improves the standard MC estimator in terms of the computational cost if the squared correlation is large enough or the coarse model is cheap.

Corollary 3.15 (Cost)**.** The cost of the single CV estimator to achieve $\mathrm{MSE}(\widehat{\mu}_L^{\mathrm{CV}}) \leq \varepsilon^2$ is

$$\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \frac{\sigma_L^2}{\varepsilon^2}(1 - \rho_{L,L-1}^2)(w_L + w_{L-1}).$$

Furthermore, the cost ratio w.r.t. the MC estimator satisfies

$$\frac{\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{CV}}\big]}{\mathbb{W}[\widehat{\mu}_L^{\mathrm{MC}}]} = (1 - \rho_{L,L-1}^2)\left(1 + \frac{w_{L-1}}{w_L}\right).$$

The CV estimator is an improvement compared to the MC estimator if the condition $w_{L-1} < \rho_{L,L-1}^2(w_{L-1} + w_L)$ is satisfied.

Proof. The result follows from Lemma 3.9 and Lemma 3.14. □

Multiple Control Variates. The results for a single CV estimator easily generalize to multiple control variates. The basic idea is to form a linear combination of the coarse models $Z_1, \ldots, Z_{L-1}$.

Definition 3.16 (Multiple CV estimator)**.** For $m \in \mathbb{N}$ and $\beta \in \mathbb{R}^{L-1}$ the *(multiple) CV estimator* is

$$\widehat{\mu}_L^{\mathrm{CV}} := \frac{1}{m}\sum_{i=1}^{m} Z_L^i - \sum_{\ell=1}^{L-1} \beta_\ell \left(\frac{1}{m}\sum_{i=1}^{m} Z_\ell^i - \mu_\ell\right).$$

We call the random variables $Z_1, \ldots, Z_{L-1}$ *control variates.* ◇

The properties of this estimator are similar to the single CV estimator in Lemma 3.14. We again want to choose β such that the variance is minimized, which yields a quadratic minimization problem. The variance minimizer is again independent of the number of samples.

Lemma 3.17 (Bias, variance)**.** Let $I := \{1, \ldots, L-1\}$ be the index set of the control variates. The CV estimator is an unbiased estimator for μ_L with variance

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \frac{\sigma_L^2 - 2C_{L,I}\beta + \beta^T C_{I,I}\beta}{m}. \tag{3.16}$$

For positive definite $C_{I,I}$ the unique minimizer for the variance is

$$\beta^* = C_{I,I}^{-1} C_{I,L}. \tag{3.17}$$

The variance for this minimizer is

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \frac{1}{m}\big(\sigma_L^2 - C_{L,I}C_{I,I}^{-1}C_{I,L}\big). \tag{3.18}$$

Proof. The proof is a straightforward extension of the proof of Lemma 3.14 to multiple dimensions. □

Let us comment on the assumption that $C_{I,I}$ is positive definite and thus invertible. This is necessary for the uniqueness of a minimizer β^*. If $C_{I,I}$ is not positive definite then Lemma 2.16 shows there exist coefficients $\widetilde{\beta} \in \mathbb{R}^{L-1}$ such that

$$\sum_{\ell=1}^{L-1} \widetilde{\beta}_\ell (Z_\ell - \mu_\ell) = 0.$$

Therefore we may w.l.o.g. remove the last control variate Z_{L-1} from the CV estimator

$$\widehat{\mu}_L^{\mathrm{CV}} := \frac{1}{m}\sum_{i=1}^{m} Z_L^i - \sum_{\ell=1}^{L-2} \beta_\ell \left(\frac{1}{m}\sum_{i=1}^{m} Z_\ell^i - \mu_\ell \right).$$

We repeat this process until the remaining control variates $I' \subsetneq \{1, \dots, L-1\}$ satisfy that $C_{I',I'}$ is positive definite or $I = \emptyset$. A minimizer β^* for the original problem is thus

$$\beta_\ell^* = \begin{cases} e_\ell^T C_{I',I'}^{-1} C_{I',L}, & \text{if } \ell \in I', \\ 0, & \text{if } \ell \notin I'. \end{cases}$$

In this case the minimizer may not be unique, since linear dependence allows us to either remove at least two different control variates from I or $Z_\ell = \mu_\ell$ almost surely. In the latter case β_ℓ^* may be chosen arbitrarily.
Let us comment on the expression for the minimal variance (3.18). We write down the covariance matrix C in block form

$$C = \begin{pmatrix} \sigma_L^2 & C_{L,I} \\ C_{I,L} & C_{I,I} \end{pmatrix}.$$

Then (3.18) without the factor $1/m$ is the *Schur complement* [61, Chapter 3] of the block $C_{I,I}$ of the covariance matrix

$$C/C_{I,I} := \sigma_L^2 - C_{L,I} C_{I,I}^{-1} C_{I,L}.$$

The multiple CV estimator reduced the variance compared to the MC estimator if the Schur complement is small enough and the coarse models are cheap. We obtain a result that is similar to Corollary 3.15.

Corollary 3.18 (Cost)**.** Let $I := \{1, \dots, L-1\}$ and $C_{I,I}$ be positive definite. Then the cost to achieve $\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] \leq \varepsilon^2$ is

$$\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \frac{1}{\varepsilon^2}(C/C_{I,I}) \sum_{\ell=1}^{L} w_\ell.$$

Furthermore, the cost ratio w.r.t. the MC estimator is

$$\frac{\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{CV}}\big]}{\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{MC}}\big]} = \left(\frac{C/C_{I,I}}{\sigma_L^2} \right) \sum_{\ell=1}^{L} \frac{w_\ell}{w_L}.$$

Proof. The corollary follows from Lemma 3.17. □

Lower variance bound. Let us now assume that we have an infinite amount of samples available for $Z_1, \dots, Z_{L-1}$ but only m samples of Z_L. This fits in the setting of control variates since by the strong law of large numbers [79, Section 5.3] we may assume that $\mu_1, \dots, \mu_{L-1}$ are known. Formally, for $n \to +\infty$ we have convergence in the almost sure sense of the sample average to the mean

$$\frac{1}{n} \sum_{i=1}^{n} Z_\ell^i \to \mu_\ell.$$

This situation occurs in practice if we have a time limit for an expensive but fast computer allowing us to sample only m times from the high fidelity model Z_L. Later we may sample more often the coarser models $Z_1, \dots, Z_{L-1}$ on a slower computer such that the error for the approximation of $\mu_1, \dots, \mu_{L-1}$ is negligible. Another situation where this occurs is if the samples of Z_L are expensive real life experiments and samples of $Z_1, \dots, Z_{L-1}$ are simulations by a computer. We first assume that we only have a single sample $m = 1$ of the high fidelity model.

Definition 3.19 (Lower variance bound of an estimator)**.** We define the *lower variance bound* for an estimator $\widehat{\mu}$ with a single evaluation $m_L := 1$ of Z_L as

$$\mathbb{V}^{\min}[\widehat{\mu}] := \lim_{m_1,\dots,m_{L-1} \to +\infty} \mathbb{V}[\widehat{\mu}],$$

where $m_1, \dots, m_{L-1}$ denotes the number of i.i.d. realizations of $Z_1, \dots, Z_{L-1}$. ◇

The restriction to a single high fidelity sample $m_L = 1$ is often not severe, since the estimators in this thesis have the property that the lower variance bound for arbitrary m_L translates to $\mathbb{V}^{\min}[\widehat{\mu}]/m_L$.
The MC estimator is special in the sense that it does not use any coarse grid samples. Therefore, the lower variance bound for MC is the variance of the high fidelity model

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\mathrm{MC}}] = \sigma_L^2.$$

We use (3.18) to conclude that the Schur complement is the lower variance bound for the CV estimators such that with $I := \{1, \dots, L-1\}$ and $J := I \cup \{L\}$

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \sigma_L^2 - C_{L,I} C_{I,I}^{-1} C_{I,L} = C_{J,J}/C_{I,I}.$$

In Chapter 4 we formally prove that $\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big]$ is the minimal variance that is achievable by any linear unbiased estimator of μ_L with only a single sample of Z_L and an infinite amount of samples of $Z_1, \dots, Z_{L-1}$.

Definition 3.20 (Lower variance bound)**.** Let $I \subseteq \{1, \dots, L-1\}$ such that $C_{I,I}$ is positive definite. Then for $J := I \cup \{L\}$ we define the *lower variance bound* or *maximally achievable variance reduction* as the Schur complement

$$\mathbb{V}_I^{\min} := \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}\big] = \sigma_L^2 - C_{L,I} C_{I,I}^{-1} C_{I,L} = C_{J,J}/C_{I,I},$$

where the CV estimator $\widehat{\mu}_L^{\mathrm{CV}}$ uses a single sample $m = 1$ and control variates $Z_{I_1}, \dots, Z_{I_{|I|}}$.
◇

The CV estimator should achieve a larger variance reduction with more control variates. Every additional model Z_ℓ allows us to choose the corresponding parameter $\beta_\ell \neq 0$. Therefore, the new estimator has more degrees of freedom which translates to an equal or smaller variance. The lower variance bounds $\mathbb{V}^{\min}$ are therefore non–increasing if we include more and more models.

Lemma 3.21 (Lower variance bound ordering)**.** For $I \subseteq \{1, \dots, L-1\}$ and $J \subseteq I$ the lower variance bound satisfies

$$\mathbb{V}_I^{\min} \leq \mathbb{V}_J^{\min}. \tag{3.19}$$

Proof. Let us denote the CV estimator with control variates indices in I with $\widehat{\mu}_L^{\mathrm{CV}}[I]$ with coefficients β. We do the same with for J with $\widehat{\mu}_L^{\mathrm{CV}}[J]$ with coefficients α. We use (3.16) and use the optimality of the coefficients to conclude

$$\begin{aligned}
\mathbb{V}_I^{\min} &= \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}[I]\big] = \min_{\beta \in \mathbb{R}^{|I|}} \sigma_L^2 - 2C_{L,I}\beta + \beta^T C_{I,I}\beta \\
\mathbb{V}_J^{\min} &= \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{CV}}[J]\big] = \min_{\alpha \in \mathbb{R}^{|J|}} \sigma_L^2 - 2C_{L,J}\alpha + \alpha^T C_{J,J}\alpha.
\end{aligned}$$

It is crucial to observe that $J \subseteq I$ and thus the first optimization problem simply minimizes over more coefficients. We fix $\beta_\ell = 0$ for all $\ell \in I \setminus J$ to obtain the second optimization problem

$$\begin{aligned}
\min_{\beta \in \mathbb{R}^{|I|}} \sigma_L^2 - 2C_{L,I}\beta + \beta^T C_{I,I}\beta \leq & \min_{\substack{\beta \in \mathbb{R}^{|I|}, \\ \beta_\ell = 0 \text{ if } \ell \in I \setminus J}} \sigma_L^2 - 2C_{L,I}\beta + \beta^T C_{I,I}\beta \\
&= \min_{\alpha \in \mathbb{R}^{|J|}} \sigma_L^2 - 2C_{L,J}\alpha + \alpha^T C_{J,J}\alpha
\end{aligned}$$

and the result (3.19) now follows. □

It is possible to extend the notion of a lower variance bound for multiple models, even if they are not high fidelity.

Remark 3.22 (Lower variance bound for multiple models)**.** We define the set of models $Q \subseteq \{1, \dots, L\}$ as the models that are evaluated once and $I \subseteq \{1, \dots, L\}$ as the models that we evaluate infinitely often. Then with $J := Q \cup I$ the covariance of the vector of estimators $(\widehat{\mu}_\ell^{\mathrm{CV}})_{\ell \in Q}$ with correlated samples satisfies

$$\mathbb{C}\mathrm{ov}^{\min} := \mathbb{C}\mathrm{ov}\big[(\widehat{\mu}_\ell^{\mathrm{CV}})_{\ell \in Q}, (\widehat{\mu}_\ell^{\mathrm{CV}})_{\ell \in Q}\big] = C_{J,J}/C_{I,I} \in \mathbb{R}^{|Q| \times |Q|}. \qquad \diamond$$

Since $\mathbb{C}\mathrm{ov}^{\min}$ is however a matrix and not single number, it is not clear what a lower bound in this context means. It is possible to reduce this (often) positive definite matrix to a single number via the trace($\mathbb{C}\mathrm{ov}^{\min}$) or by looking at a specific linear combination $\beta^T \mathbb{C}\mathrm{ov}^{\min}\beta$.

Problems of the CV estimator. The CV estimator requires us to know the expectation of the control variates $\mu_1, \dots, \mu_{L-1}$. These values are often unavailable and we need to estimate them. Therefore, we could again devise a CV estimator.

Example 3.23 (Three level nested CV estimator)**.** For the purpose of this example define the CV estimator for μ_1 as the MC estimator

$$\widehat{\mu}_1^{\mathrm{CV}} := \frac{1}{m_1} \sum_{i=1}^{m_1} Z_1^i.$$

We then define an approximation for the CV estimator for μ_2 such that

$$\widehat{\mu}_2^{\mathrm{CV}} = \frac{1}{m_2}\sum_{i=1}^{m_2} Z_2^i - \beta_1\left(\frac{1}{m_2}\sum_{i=1}^{m_2}(Z_1^i - \widehat{\mu}_1^{\mathrm{CV}})\right).$$

Now we continue with the third level

$$\widehat{\mu}_3^{\mathrm{CV}} = \frac{1}{m_3}\sum_{i=1}^{m_3} Z_3^i - \sum_{\ell=1}^{2}\beta_\ell\left(\frac{1}{m_3}\sum_{i=1}^{m_3}(Z_\ell^i - \widehat{\mu}_\ell^{\mathrm{CV}})\right). \qquad \diamond$$

The example shows that there are a several degrees of freedom and choices to make.

- We may choose the coefficients β for both estimators $\widehat{\mu}_2^{\mathrm{CV}}$ and $\widehat{\mu}_3^{\mathrm{CV}}$ to dependent on each other.
- We want to distribute the number of samples m_1, m_2, m_3 to achieve a small variance.
- It is unclear how to choose the independence structure of the random variables, i.e. it might make sense to use independent samples to estimate $\widehat{\mu}_1^{\mathrm{CV}}$ and $\widehat{\mu}_2^{\mathrm{CV}}$, but not for $\widehat{\mu}_3^{\mathrm{CV}}$.
- It is trivial to generalize this estimator for more levels, however, the optimality of this approach is not clear.

The Multifidelity Monte Carlo method in Section 3.4 and Approximate Control Variate method in Section 3.5 both estimate μ_L in similar fashion shown in Example 3.23.

3.4 Multifidelity Monte Carlo

We continue with the Multifidelity Monte Carlo (MFMC) estimator in [107]. This method assumes that the mean values $\mu_1, \dots, \mu_{L-1}$ of the control variates are unknown and have to be estimated. The MFMC estimator is a CV estimator with a specific number of samples m and independence of the realizations.

Definition 3.24 (MFMC estimator)**.** For $m_1, \dots, m_L \in \mathbb{N}$ and $\beta \in \mathbb{R}^{L-1}$ we define the *MFMC estimator*

$$\widehat{\mu}_L^{\mathrm{MFMC}} := \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=1}^{L-1}\beta_\ell\left(\frac{1}{m_{\ell+1}}\sum_{i=1}^{m_{\ell+1}} Z_\ell^i - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^i\right). \tag{3.20}$$

$\diamond$

Similar to the CV estimator we want to choose β and m such that the variance of the estimator is minimized. As it turns out, the optimal choice of β is independent of the choice of m allowing us to optimize them separately. The next lemma optimizes the β and is a summary of the results of Section 3 from [107]. We simplify the notation and abbreviate $\rho_{L,0} := 0$.

Lemma 3.25 (Bias, variance [107, Section 3])**.** Let $m_1 \geq \cdots \geq m_L$ and assume $\sigma_\ell^2 > 0$ for all $\ell \in \{1, \dots, L\}$. Then the MFMC estimator is an unbiased estimator for μ_L with variance

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] = \frac{\sigma_L^2}{m_L} + \sum_{\ell=1}^{L-1} \left(\frac{1}{m_{\ell+1}} - \frac{1}{m_\ell} \right) (-2\beta_\ell C_{L\ell} + \beta_\ell^2 \sigma_\ell^2). \tag{3.21}$$

A variance minimizer $\beta^* \in \mathbb{R}^{L-1}$ is

$$\beta_\ell^* = C_{L\ell}/\sigma_\ell^2 \quad \text{for all } \ell \in \{1, \dots, L-1\} \tag{3.22}$$

and the minimal variance satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] = \sigma_L^2 \sum_{\ell=1}^{L} \frac{\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2}{m_\ell}. \tag{3.23}$$

Proof. The proof is scattered throughout Section 3 of [107]. The unbiasedness follows directly from the linearity of the expectation. The result in (3.21) is a straightforward calculation done in [107, Lemma 3.2, Lemma 3.3]. A step in the proof is to compute

$$\mathbb{C}\mathrm{ov}\left[\frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_\ell^i, \frac{1}{m_j} \sum_{i=1}^{m_j} Z_j^i \right] = \frac{1}{\max\{m_\ell, m_j\}} C_{\ell,j}.$$

The assumption $m_1 \geq \cdots \geq m_L$ is then used to obtain the easier manageable form (3.21). The proof of (3.22) is given in [107, Theorem 3.4]. We use $m_1 \geq \cdots \geq m_L$ to ensure that the expression

$$\left(\frac{1}{m_{\ell+1}} - \frac{1}{m_\ell} \right) \geq 0$$

is non–negative for all $\ell \in \{1, \dots, L-1\}$. W.l.o.g. we may assume that this expression is larger than zero, otherwise the value of β_ℓ does not influence the variance in (3.21) and we may choose $\beta_\ell^* = C_{L\ell}/\sigma_\ell^2$ as in (3.22). For $\sigma_\ell^2 > 0$ equation (3.21) as function of β is a sum of $L-1$ parabolas with positive leading factor. We separately minimize them to obtain the unique minimizer

$$\beta_\ell^* = C_{L\ell}/\sigma_\ell^2 \quad \text{for all } \ell \in \{1, \dots, L-1\}.$$

Inserting this optimal β^* into (3.21) then leads to (3.23). □

The condition $m_1 \geq \cdots \geq m_L$ is often satisfied for hierarchical models, where we require few evaluations of the high fidelity model and allow for more evaluations of the coarser models. Furthermore, $\sigma_\ell^2 > 0$ is a mild assumption, since $\sigma_\ell^2 = 0$ implies that Z_ℓ is almost surely constant and thus $C_{L\ell} = 0$. The corresponding term in (3.21) is then zero and we remove Z_ℓ from the estimator. The minimizer β^* is unique if the stricter inequality $m_1 > \cdots > m_L$ is satisfied. This is desirable since for $m_{j+1} = m_j$ the terms associated with the control variate Z_j are zero and we remove this model

$$\begin{aligned} \widehat{\mu}_L^{\mathrm{MFMC}} &= \frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=1}^{L-1} \beta_\ell \left(\frac{1}{m_{\ell+1}} \sum_{i=1}^{m_{\ell+1}} Z_\ell^i - \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_\ell^i \right) \\ &= \frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^i - \sum_{\substack{\ell=1, \\ \ell \neq j}}^{L-1} \beta_\ell \left(\frac{1}{m_{\ell+1}} \sum_{i=1}^{m_{\ell+1}} Z_\ell^i - \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_\ell^i \right). \end{aligned}$$

Sample allocation. We now answer the question how to optimally allocate the number of samples m_ℓ for every level ℓ. The goal is to minimize the variance not exceeding a fixed computational budget $\mathbb{W}^{\text{budget}} > 0$. The MFMC estimator requires m_ℓ samples of Z_ℓ and each sample incurs costs of $w_\ell > 0$. This is formalized in the following relaxed optimization problem, where we allow fractional samples:

$$\begin{aligned} \min_{m_1,\dots,m_L \in \mathbb{R}} J(m) := \mathbb{V}\big[\widehat{\mu}_L^{\text{MFMC}}\big] &= \sigma_L^2 \sum_{\ell=1}^{L} \frac{\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2}{m_\ell} \\ \text{such that} \qquad \sum_{\ell=1}^{L} m_\ell w_\ell &= \mathbb{W}^{\text{budget}}, \\ m_\ell &\geq m_{\ell+1} \qquad \text{for all } \ell \in \{1,\dots,L-1\}, \\ m_L &\geq 0. \end{aligned} \tag{3.24}$$

We use the ordering constraints to apply (3.23) of Lemma 3.25 which ensures that the expression for $\mathbb{V}\big[\widehat{\mu}_L^{\text{MFMC}}\big]$ is valid. The budget constraint

$$\sum_{\ell=1}^{L} m_\ell w_\ell \leq \mathbb{W}^{\text{budget}}$$

is satisfied with equality at a minimizer, otherwise we could linearly scale up $m_1,\dots,m_L$ to further reduce the variance. We include this constraint as equality constraint, since these are easier to handle than inequality constraints in convex optimization.

Theorem 3.26 (Optimal sample allocation [107, Theorem 3.4, Corollary 3.5])**.**
Let $\sigma_\ell^2 > 0$ for $\ell \in \{1,\dots,L\}$ and assume the ordering

$$\rho_{L,1}^2 < \cdots < \rho_{L,L}^2. \tag{3.25}$$

Furthermore, let the costs and correlations satisfy

$$w_{\ell+1}(\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2) > w_\ell(\rho_{L,\ell+1}^2 - \rho_{L,\ell}^2) \tag{3.26}$$

for all $\ell \in \{1,\dots,L-1\}$. Then the unique variance minimal sample allocation m^* of (3.24) satisfies

$$m_\ell^* = \frac{\mathbb{W}^{\text{budget}}}{\sum_{j=1}^{L}\big(w_j(\rho_{L,j}^2 - \rho_{L,j-1}^2)\big)^{1/2}} \left(\frac{\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2}{w_\ell}\right)^{1/2} \tag{3.27}$$

for all $\ell \in \{1,\dots,L\}$. The variance for this optimal sample allocation satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\text{MFMC}}\big] = \frac{\sigma_L^2}{\mathbb{W}^{\text{budget}}} \left(\sum_{\ell=1}^{L}\big(w_\ell(\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2)\big)^{1/2}\right)^2. \tag{3.28}$$

Proof. A detailed proof is given in [107, Theorem 3.4]. The assumptions (3.25) and (3.26) are used to ensure that the constraint $m_1 \geq \cdots \geq m_L$ is satisfied with strict inequality at the optimum m^*. Then (3.24) is minimized using basic convex optimization. □

The variance $\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big]$ is inversely proportional to $\mathbb{W}^{\mathrm{budget}}$. This allows us to compare the cost to the MC estimator similar to Corollary 3.18 for the CV estimator. The main difference is that the variance reduction of the MFMC estimator only depends on the correlations between the high fidelity model and the low fidelity models $\rho_{L,1}, \dots, \rho_{L,L-1}$ but no correlation between two low fidelity models. In contrast, the CV estimator depends on the Schur complement of C, which takes other correlations into account.

Lemma 3.27 (Cost)**.** Let the assumptions of Theorem 3.26 be true. Then the cost to achieve $\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] \leq \varepsilon^2$ is

$$\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] = \frac{\sigma_L^2}{\varepsilon^2}\left(\sum_{\ell=1}^{L}\big(w_\ell(\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2)\big)^{1/2}\right)^2. \tag{3.29}$$

Moreover, the cost compared to the MC estimator satisfies

$$\frac{\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big]}{\mathbb{W}[\widehat{\mu}_L^{\mathrm{MC}}]} = \left(\sum_{\ell=1}^{L}\left(\frac{w_\ell}{w_L}(\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2)\right)^{1/2}\right)^2. \tag{3.30}$$

Proof. The proof is a straightforward application of Theorem 3.26. □

Graph based model selection. We have to devise an algorithm to ensure that the assumptions of Theorem 3.26 are satisfied. Furthermore, we want to select or order the models in such a way that the cost of the estimator is minimized. This problem is called model selection problem.

Definition 3.28 (Model selection problem)**.** The *model selection problem* for MFMC is a minimization problem

$$\min_P \quad J(P) := \sum_{\ell=1}^{|P|}\Big(w_{P_\ell}(\rho_{L,P_\ell}^2 - \rho_{L,P_{\ell-1}}^2)\Big)^{1/2}, \tag{3.31}$$

where $P = (P_1, P_2, \dots, P_{|P|-1}, L)^T$ is a vector of length $|P| \leq L$ with $P_\ell \in \{1, \dots, L\}$. We further require that P satisfies

$$\sigma_{P_\ell}^2 > 0 \qquad \text{for all } \ell \in \{1, \dots, |P|\}, \tag{3.32}$$

$$\rho_{L,P_1}^2 < \dots < \rho_{L,P_{|P|}}^2, \tag{3.33}$$

$$w_{P_{\ell+1}}(\rho_{L,P_\ell}^2 - \rho_{L,P_{\ell-1}}^2) > w_{P_\ell}(\rho_{L,P_{\ell+1}}^2 - \rho_{L,P_\ell}^2) \qquad \text{for all } \ell \in \{1, \dots, |P|-1\}. \tag{3.34}$$

We define $\rho_{L,P_0}^2 := 0$. ◇

The conditions (3.32), (3.33) and (3.34) are used to apply Lemma 3.27. The cost function J in (3.31) is up to a constant equal to the square root of the cost of the MFMC estimator in (3.29). The model selection $P = (L)^T$ corresponds to the MC estimator and thus the minimizer in (3.31) is guaranteed to lead to an estimator with variance not exceeding MC. Without model selection this property is in general not true.
The algorithm originally proposed by Peherstorfer [107, Algorithm 1] applies a brute force strategy and checks all valid P that satisfy (3.32), (3.33), (3.34). The model selection P with the smallest variance is then used. This strategy has costs exponential in the number of models L.

We propose a graph based approach where we reformulate (3.31) as a shortest path problem. To explain the main idea we first neglect the constraints. Let Z_0 be an artificial model with $w_0 := 0$, $\mathbb{C}\mathrm{ov}[Z_L, Z_0] := 0$, $\sigma_0^2 > 0$ and define edge weights

$$q_{ij} := \big(w_i(\rho_{L,i}^2 - \rho_{L,j}^2)\big)^{1/2}.$$

We then define the directed graph G with vertices V, edges E and edge weights Q

$$G := (V, E, Q), \quad V := \{0, \dots, L\}, \quad E := V \times V, \quad Q := (q_{ij})_{i,j=0}^{L}.$$

The cost of a path P from 0 to L is exactly $J(P)$ in (3.31). Furthermore, every path corresponds to a valid model selection and vice versa. Solving the model selection problem is thus equivalent to finding a shortest path in a directed graph.
Let us now introduce the constraints that are all of local nature. First, if (3.32) is not satisfied for some model Z_ℓ, then we do not add any edges to the vertex ℓ. This ensures that no path from 0 to L uses this model. Condition (3.33) is satisfied for every path P if we only add edges (i, j) if $\rho_{L,i}^2 < \rho_{L,j}^2$. The constraint (3.34) is tricky since it is a condition between three models and thus three vertices. We solve this issue by introducing another dimension in the vertex set of the graph such that nodes are now pairs (i, j). We then introduce edges of the form $((i, j), (j, k))$ if condition (3.34) is satisfied. A path which uses this edge means that the models i, j were selected and k is the next model which incurs costs of q_{jk}. Let us formally define this graph G.

$$\begin{aligned}
G &:= (V, E, Q),\\
V &:= \{0, \dots, L\} \times \{0, \dots, L\},\\
E &:= \Big\{((i,j),(j,k)) \in V \times V \,|\, i \neq j, i \neq k, j \neq k \text{ and}\\
&\qquad \sigma_i^2 > 0,\ \sigma_j^2 > 0,\ \sigma_k^2 > 0, \text{ and}\\
&\qquad \rho_{L,j}^2 < \rho_{L,k}^2, \text{ and}\\
&\qquad w_k(\rho_{L,j}^2 - \rho_{L,i}^2) > w_j(\rho_{L,k}^2 - \rho_{L,j}^2)\Big\}\\
&\qquad \cup \{((0,0),(0,\ell)) \,|\, \ell \in \{1, \dots, L\}\}\\
&\qquad \cup \{((\ell, L),(L, L)) \,|\, \ell \in \{1, \dots, L\}\},\\
Q &:= (q_{ijjk})_{i,j,k=0}^{L}, \quad q_{ijjk} := q_{jk}.
\end{aligned} \tag{3.35}$$

The goal is to find a shortest path from $(0, 0)$ to (L, L).

Lemma 3.29 (Shortest path model selection)**.** Finding the shortest path from $(0, 0)$ to (L, L) in (3.35) is equivalent to the model selection problem in Definition 3.28.

Proof. We first show that every path P from $(0, 0)$ to (L, L) in G delivers a valid model selection and vice versa. For some $K \in \{1, \dots, L\}$ the path P is

$$P = ((0,0),(0,\ell_1),(\ell_1,\ell_2),(\ell_2,\ell_3),\dots,(\ell_K, L),(L, L)) \tag{3.36}$$

and from this define the reduced path or model selection

$$P^{\mathrm{red}} := (\ell_1, \ell_2, \ell_3, \dots, \ell_K, L). \tag{3.37}$$

The reduced path P^{red} satisfies (3.32), (3.33) and (3.34) by construction. We only verify this for last condition. For $\ell \in \{1, \dots, K\}$ and with $P_0^{\mathrm{red}} := 0$ abbreviate the vertices $i := P_{\ell-1}^{\mathrm{red}}$, $j := P_\ell^{\mathrm{red}}$ and $k := P_{\ell+1}^{\mathrm{red}}$. The construction of P^{red} from P shows

$$P = (\dots, (i, j), (j, k), \dots)$$

and thus $((i,j),(j,k)) \in E$. The definition of the edge set E in (3.35) shows (3.32)

$$w_k(\rho_{L,j}^2 - \rho_{L,i}^2) > w_j(\rho_{L,k}^2 - \rho_{L,j}^2).$$

The verification of the other conditions (3.32) and (3.33) is similar. We conclude that P^{red} is a valid model selection. For the other direction let P^{red} be a valid model selection given as in (3.37). From this we define P as in (3.36). It is now straightforward to verify that P is actually a path in G from the assumption that P^{red} is a valid model selection satisfying (3.32), (3.33) and (3.34).
All that remains is to show that the cost of a path P is equal to the cost of the respective model selection. We denote the first entry of $P_\ell = (i,j)$ with $(P_\ell)_1 := i$. The definition of q in (3.35) shows that the cost satisfies

$$\sum_{\ell=2}^{|P|} q_{(P_{\ell-1})_1 (P_{\ell-1})_2 (P_\ell)_1 (P_\ell)_2} = \sum_{\ell=2}^{|P|-1} q_{(P_\ell)_1 (P_\ell)_2},$$

where we used that the last term is $q_{L,L} = 0$. The definition of P^{red} in (3.37) shows

$$\sum_{\ell=2}^{|P|-1} q_{(P_\ell)_1 (P_\ell)_2} = \sum_{\ell=1}^{|P^{\mathrm{red}}|} q_{P_{\ell-1}^{\mathrm{red}} P_\ell^{\mathrm{red}}} = \sum_{\ell=1}^{|P^{\mathrm{red}}|} \left(w_{P_\ell^{\mathrm{red}}} (\rho_{L,P_\ell^{\mathrm{red}}}^2 - \rho_{L,P_{\ell-1}^{\mathrm{red}}}^2) \right)^{1/2} = J(P^{\mathrm{red}}),$$

which is exactly the cost in (3.31). □

We remark that computing a shortest path in a directed graph G with non–negative weights can be done efficiently in the sense that the costs are polynomial in the number of models L. A well–known algorithm to compute shortest paths is Dijkstra's algorithm, see [13, Section 1.5] or [2, Section 7.4].

Asymptotic analysis. Let us return to the setting that we know asymptotic information of the models Z_ℓ. We assume that the cost w_ℓ increases at most geometrically and the difference $\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2$ decreases sufficiently fast. The latter is achieved if the difference of the variance of two consecutive levels converges to zero at a certain rate. The overall variance of the estimator is then given in terms of an asymptotic expansion. We summarize the result of [106, Lemma 2] in the next lemma.

Lemma 3.30 (Asymptotic variance)**.** Let the assumptions of Theorem 3.26 be true and for all $\ell \in \mathbb{N}$

$$\mathrm{Bias}(Z_\ell) \leq c 2^{-\gamma_{\mathrm{Bias}} \ell}, \tag{3.38}$$

$$\mathbb{V}[Z_\ell - Z_{\ell-1}] \leq c 2^{-\gamma_{\mathrm{Var}} \ell}, \tag{3.39}$$

$$w_\ell \leq c 2^{\gamma_{\mathrm{Cost}} \ell}. \tag{3.40}$$

Furthermore, assume uniform lower and upper bounds on the variance for all $\ell \in \mathbb{N}$

$$0 < \sigma_-^2 \leq \sigma_\ell^2 \leq \sigma_+^2 < +\infty. \tag{3.41}$$

Then the bias for the MFMC estimator is

$$\mathrm{Bias}(\widehat{\mu}_L^{\mathrm{MFMC}}) \leq c 2^{-\gamma_{\mathrm{Bias}} L} \tag{3.42}$$

and the variance satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] \leq \frac{c}{\mathbb{W}^{\mathrm{budget}}} \left(\sum_{\ell=1}^{L} 2^{(\gamma_{\mathrm{Cost}} - \gamma_{\mathrm{Var}})\ell/2} \right)^2. \tag{3.43}$$

Proof. The MFMC estimator $\widehat{\mu}_L^{\mathrm{MFMC}}$ is an unbiased estimator for μ_L and thus (3.38) shows (3.42). The proof for the variance (3.43) is given in [106, Lemma 1, Lemma 2] and we only write down the main idea. We have to bound $\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2$ in (3.28)

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] = \frac{\sigma_L^2}{\mathbb{W}^{\mathrm{budget}}} \left(\sum_{\ell=1}^{L} \big(w_\ell (\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2) \big)^{1/2} \right)^2 . \tag{3.44}$$

We assume $0 < \rho_{L,1} < \cdots < \rho_{L,L}$, otherwise redefine some QoIs from Z_ℓ to $-Z_\ell$. Combine this with $\rho_{L,\ell} \in [0,1]$ to obtain the bound

$$\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2 = (\rho_{L,\ell} + \rho_{L,\ell-1})(\rho_{L,\ell} - \rho_{L,\ell-1}) \leq 2(\rho_{L,\ell} - \rho_{L,\ell-1}).$$

W.l.o.g. we assume $\sigma_\ell^2 = 1$ for all $\ell \in \{1, \dots, L\}$ due to (3.41) and apply Höelder's inequality to get

$$\begin{aligned} \rho_{L,\ell} - \rho_{L,\ell-1} &= \mathbb{C}\mathrm{ov}[Z_L, Z_\ell - Z_{\ell-1}] \\ &= \mathbb{C}\mathrm{ov}[Z_L - Z_{\ell-1}, Z_\ell - Z_{\ell-1}] + \mathbb{C}\mathrm{ov}[Z_{\ell-1}, Z_\ell - Z_{\ell-1}] \\ &\leq \big(\mathbb{V}[Z_L - Z_{\ell-1}] \mathbb{V}[Z_\ell - Z_{\ell-1}] \big)^{1/2} + \mathbb{C}\mathrm{ov}[Z_{\ell-1}, Z_\ell - Z_{\ell-1}]. \end{aligned}$$

The variance assumption (3.39) is used to bound the first term by $c2^{-\gamma_{\mathrm{Var}}\ell}$. Furthermore, the second term is non–positive

$$\mathbb{C}\mathrm{ov}[Z_{\ell-1}, Z_\ell - Z_{\ell-1}] = \rho_{\ell,\ell-1} - 1 \leq 0 \tag{3.45}$$

and thus $\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2 \leq c2^{-\gamma_{\mathrm{Var}}\ell}$. We insert this and use the sample cost (3.40) to bound the overall cost in (3.44), which yields (3.43). □

The property (3.45) is crucially needed for the proof. We emphasize this, since since a direct application of the Cauchy–Schwarz inequality leads to only half the rate

$$\rho_{L,\ell} - \rho_{L,\ell-1} = \mathbb{C}\mathrm{ov}[Z_L, Z_\ell - Z_{\ell-1}] \leq \big(\mathbb{V}[Z_L] \mathbb{V}[Z_\ell - Z_{\ell-1}] \big)^{1/2} \leq c2^{-\gamma_{\mathrm{Var}}\ell/2}.$$

Let us summarize the asymptotic complexity of the estimator in the next theorem. The proof is very similar to the proof of [56, Theorem 1] and [31, Theorem 1]. The statement is a slight generalization of the result in [106, Theorem 1, Corollary 1] since it takes rounding into account.

Theorem 3.31 (Asymptotic cost [106, Theorem 1, Corollary 1])**.** Let the assumptions of Lemma 3.30 be true. Then for all $\varepsilon \in (0, 1/e]$ there exists a fine level L and a number of samples $m_1, \dots, m_L$ such that cost to achieve $\mathrm{MSE}(\widehat{\mu}_L^{\mathrm{MFMC}}) \leq \varepsilon^2$ is bounded

$$\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] \leq c\varepsilon^{-\gamma_{\mathrm{Cost}}/\gamma_{\mathrm{Bias}}} + c \begin{cases} \varepsilon^{-2}, & \text{if } \gamma_{\mathrm{Var}} > \gamma_{\mathrm{Cost}}, \\ \varepsilon^{-2} \log(\varepsilon)^2, & \text{if } \gamma_{\mathrm{Var}} = \gamma_{\mathrm{Cost}}, \\ \varepsilon^{-2 - \frac{\gamma_{\mathrm{Cost}} - \gamma_{\mathrm{Var}}}{\gamma_{\mathrm{Bias}}}}, & \text{if } \gamma_{\mathrm{Var}} < \gamma_{\mathrm{Cost}}. \end{cases} \tag{3.46}$$

Proof. Fix $\varepsilon^2 > 0$ and choose

$$L := -\log(\varepsilon)/\gamma_{\mathrm{Bias}} + L_0 \tag{3.47}$$

for a suitable $L_0 \leq c$ large enough such that $L \in \mathbb{N}$ and that the result (3.42) gives

$$\mathrm{Bias}(\widehat{\mu}_L^{\mathrm{MFMC}})^2 \leq c2^{-2\gamma_{\mathrm{Bias}}L} = c2^{-2\gamma_{\mathrm{Bias}}L_0}\varepsilon^2 \leq \varepsilon^2/2.$$

We further choose

$$\mathbb{W}^{\text{budget}} = c\varepsilon^{-2}\left(\sum_{\ell=1}^{L} 2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})\ell/2}\right)^2/2$$

and combine these results with a bias-variance decomposition and (3.43)

$$\begin{aligned}\text{MSE}(\widehat{\mu}_L^{\text{MFMC}}) &= \text{Bias}(\widehat{\mu}_L^{\text{MFMC}})^2 + \mathbb{V}\big[\widehat{\mu}_L^{\text{MFMC}}\big] \\ &\leq \text{Bias}(\widehat{\mu}_L^{\text{MFMC}})^2 + \frac{c}{\mathbb{W}^{\text{budget}}}\left(\sum_{\ell=1}^{L} 2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})\ell/2}\right)^2 \\ &\leq \varepsilon^2.\end{aligned}$$

The asymptotic expression for the cost $\mathbb{W}\big[\widehat{\mu}_L^{\text{MFMC}}\big] = \mathbb{W}^{\text{budget}}$ without ceiling the number of samples m follows by looking at the three distinct cases. First, for $\gamma_{\text{Var}} > \gamma_{\text{Cost}}$ we use the properties of the geometric sum to obtain

$$\sum_{\ell=1}^{L} 2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})\ell/2} \leq \sum_{\ell=1}^{\infty} 2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})\ell/2} \leq c,$$

where we used the fact that the exponent is negative. We thus obtain $\mathbb{W}\big[\widehat{\mu}_L^{\text{MFMC}}\big] = c\varepsilon^{-2}$. Secondly, for $\gamma_{\text{Var}} = \gamma_{\text{Cost}}$ the sum is bounded by

$$\left(\sum_{\ell=1}^{L} 2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})\ell/2}\right)^2 = \left(\sum_{\ell=1}^{L} 1\right)^2 = L^2.$$

Inserting L from (3.47) adds the additional logarithmic factor $\mathbb{W}\big[\widehat{\mu}_L^{\text{MFMC}}\big] = c\varepsilon^{-2}\log(\varepsilon)^2$. Finally, for $\gamma_{\text{Var}} < \gamma_{\text{Cost}}$ the geometric sum now grows faster

$$\left(\sum_{\ell=1}^{L} 2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})\ell/2}\right)^2 \leq c\big(2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})(L+1)/2}\big)^2 = c2^{(\gamma_{\text{Cost}}-\gamma_{\text{Var}})L}.$$

Again, if we insert L from (3.47) we obtain the costs of $\mathbb{W}\big[\widehat{\mu}_L^{\text{MFMC}}\big] = c\varepsilon^{-2}\varepsilon^{-(\gamma_{\text{Cost}}-\gamma_{\text{Var}})/\gamma_{\text{Bias}}}$. Up until this point we have neglected the ceiling of the number of samples m. The geometric cost increase (3.40) ensures that the cost of ceiling is dominated by cw_L, which is a constant times the cost of a single high fidelity sample. We combine this with (3.47) to obtain

$$cw_L \leq c2^{\gamma_{\text{Cost}}L} \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}},$$

which is the additional term for the cost $\mathbb{W}\big[\widehat{\mu}_L^{\text{MFMC}}\big]$ in (3.46). □

The complexity of the MFMC estimator is significantly smaller than the complexity of the MC estimator which has the complexity expression (3.46) with $\gamma_{\text{Var}} = 0$. In particular, for $\gamma_{\text{Var}} > \gamma_{\text{Cost}}$ and assuming that the cost for rounding does not dominate, that is,

$$\gamma_{\text{Cost}}/\gamma_{\text{Bias}} \leq 2,$$

we have the optimal asymptotic cost of ε^{-2} as for the MC estimator that samples directly from Z without discretization, see Corollary 3.10. For completeness we state the asymptotic cost compared to MC.

Corollary 3.32 (Asymptotic cost quotient)**.** Let the assumptions of Theorem 3.31 be true. Then the cost of the MFMC estimator compared to the MC estimator to reach a MSE of $\varepsilon^2 \in (0, 1]$ satisfies

$$\frac{\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big]}{\mathbb{W}[\widehat{\mu}_L^{\mathrm{MC}}]} = c\varepsilon^2 + c\begin{cases} \varepsilon^{\frac{\gamma_{\mathrm{Cost}}}{\gamma_{\mathrm{Bias}}}}, & \text{if } \gamma_{\mathrm{Var}} > \gamma_{\mathrm{Cost}}, \\ \varepsilon^{\frac{\gamma_{\mathrm{Cost}}}{\gamma_{\mathrm{Bias}}}} \log(\varepsilon)^2, & \text{if } \gamma_{\mathrm{Var}} = \gamma_{\mathrm{Cost}}, \\ \varepsilon^{\frac{\gamma_{\mathrm{Var}}}{\gamma_{\mathrm{Bias}}}}, & \text{if } \gamma_{\mathrm{Var}} < \gamma_{\mathrm{Cost}}. \end{cases}$$

Proof. Combine the expressions of Theorem 3.31 and Theorem 3.11. □

We return to the PDE Example 3.12 and there the MFMC estimator achieves the optimal cost of $\mathcal{O}(\varepsilon^{-2})$ for low dimensions. These costs are asymptotically equal to the MC estimator that directly samples from Z.

Example 3.33 (PDE example)**.** Let the assumptions of Theorem 3.31 be true. Then if Z is defined as in (2.16) such that Corollary 2.41 applies, we have $\gamma_{\mathrm{Bias}} = 2$, $\gamma_{\mathrm{Var}} = 4$ and $\gamma_{\mathrm{Cost}} = d$, where d is the dimension of D. The cost is minimal for $d \leq 3$ and close to optimal for $d = 4$. The result for different dimensions is summarized in Table 3.2. ◇

Dimension d	1	2	3	4	5	6
Cost	ε^{-2}	ε^{-2}	ε^{-2}	$\varepsilon^{-2}\log(\varepsilon)^2$	$\varepsilon^{-2.5}$	ε^{-3}
Cost increase	4	4	4	≈ 4	≈ 5.66	8

Table 3.2: Cost of the MFMC estimator w.r.t. to the dimension d of the domain D. The row "Cost increase" denotes the factor by which the total cost increases if we require half the RMSE ε. The QoI is (2.16) obtained from the elliptic PDE example.

Lower variance bound. We examine the behaviour for the MFMC estimator if we increase the number of coarse grid samples $m_1, \ldots, m_{L-1}$ to infinity. We might expect the same lower variance bound $\mathbb{V}_I^{\min}$ with $I = \{1, \ldots, L-1\}$ as for the CV estimator, since $\mu_1, \ldots, \mu_{L-1}$ are known. However, as the authors of [62, Theorem 1] show, this is not the case. The MFMC estimator only reaches the bound $\mathbb{V}_{\{L-1\}}^{\min}$, which is the lower variance bound for a single CV estimator using Z_{L-1}. We provide two proofs, one were we directly increase the number of samples and one were we decrease the cost for the coarse models to zero.

Theorem 3.34 (Lower variance bound [62, Theorem 1])**.** Let the assumption of Theorem 3.26 be true. Then the lower variance bound for the MFMC estimator satisfies

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\mathrm{MFMC}}] = \mathbb{V}_{\{L-1\}}^{\min} = \sigma_L^2(1 - \rho_{L,L-1}^2). \tag{3.48}$$

Proof (sample based): Use the limit of $m_\ell \to +\infty$ for $\ell \in \{1, \ldots, L-1\}$ in (3.23) to obtain the right side of (3.48). The resulting expression is equal to (3.15). □

Proof (cost based): We have to be careful to ensure that assumption (3.26) of Theorem 3.26 is satisfied if we let $w_1, \ldots, w_{L-1}$ go to zero. Let $\alpha \to 0$ with $\alpha > 0$ and redefine the costs as $\alpha w_1, \ldots, \alpha w_{L-1}$. Now assumption (3.26) is satisfied for all $\alpha \leq 1$. We further

define $\mathbb{W}^{\text{budget}} = w_L$ and from (3.27) we conclude $m_\ell \to +\infty$ for $\ell \in \{1, \dots, L-1\}$ and $m_L \to 1$ for $\alpha \to 0$. We obtain the result from (3.28)

$$\begin{aligned}\mathbb{V}^{\min}[\widehat{\mu}_L^{\text{MFMC}}] &= \lim_{\alpha \to 0} \mathbb{V}\big[\widehat{\mu}_L^{\text{MFMC}}\big] \\ &= \lim_{\alpha \to 0} \frac{\sigma_L^2}{\mathbb{W}^{\text{budget}}} \left((w_L(\rho_{L,L}^2 - \rho_{L,L-1}^2))^{1/2} + \sum_{\ell=1}^{L-1} \big(\alpha w_\ell (\rho_{L,\ell}^2 - \rho_{L,\ell-1}^2)\big)^{1/2} \right)^2 \\ &= \sigma_L^2 (1 - \rho_{L,L-1}^2). \end{aligned}$$

□

The MFMC estimator is not optimal in the limit of infinitely many low fidelity models. From Lemma 3.21 we obtain a comparison to the full CV estimator

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\text{MFMC}}] = \mathbb{V}^{\min}_{\{L-1\}} \geq \mathbb{V}^{\min}_{\{1,\dots,L-1\}} = \mathbb{V}^{\min}[\widehat{\mu}_L^{\text{CV}}].$$

This inequality is often strict, which we later show in numerical experiments in Section 4. We want to give the reader a deeper understanding of why the variance reduction is not optimal. A close inspection of the MFMC estimator shows almost sure convergence to the single CV estimator with control variate Z_{L-1}. In particular, the correlation information between the models $Z_1, \dots, Z_{L-2}$ and Z_L is lost. The reason is that some of the sums in the definition of the MFMC estimator (3.20) almost surely converge to zero

$$\lim_{m_1,\dots,m_{L-1} \to +\infty} \left(\frac{1}{m_{\ell+1}} \sum_{i=1}^{m_{\ell+1}} Z_\ell^i - \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_\ell^i \right) = 0 \quad \text{for all } \ell \in \{1, \dots, L-2\}.$$

For $\ell \in \{1, \dots, L-2\}$ the random variables $Z_\ell^1, \dots, Z_\ell^{m_L}$ that are correlated with the samples of the fine model $Z_L^1, \dots, Z_L^{m_L}$ disappear.

Lemma 3.35 (Almost sure convergence to single CV)**.** The MFMC estimator $\widehat{\mu}_L^{\text{MFMC}}$ converges almost surely to the single CV estimator $\widehat{\mu}_L^{\text{CV}}$ with control variate Z_{L-1} and m_L samples

$$\lim_{m_1,\dots,m_{L-1} \to +\infty} \widehat{\mu}_L^{\text{MFMC}} = \widehat{\mu}_L^{\text{CV}}.$$

Proof. Let us rewrite the MFMC estimator defined in (3.20) such that the first m_L samples are all grouped together

$$\widehat{\mu}_L^{\text{MFMC}} = \sum_{i=1}^{m_L} \left(\frac{1}{m_L} Z_L^i - \sum_{\ell=1}^{L-1} \beta_\ell \left(\frac{1}{m_{\ell+1}} Z_\ell^i - \frac{1}{m_\ell} Z_\ell^i \right) \right) + \beta_{L-1} \frac{1}{m_{L-1}} \sum_{i=m_L+1}^{m_{L-1}} Z_{L-1}^i + R,$$

where R is the remainder of the estimator. If we take the limit $m_1, \dots m_{L-1}$ to infinity, the term $\frac{1}{m_\ell} Z_\ell^i$ converges almost surely to zero. This is true for all random variables in the left sum except for $\frac{1}{m_{\ell+1}} Z_\ell^i$ with $\ell = L-1$. By the strong law of large numbers R converges almost surely to zero and the remaining term converges almost surely to $\beta_{L-1}\mu_{L-1}$. We conclude almost sure convergence of the MFMC estimator to the CV estimator for $m_1, \dots, m_{L-1} \to +\infty$

$$\widehat{\mu}_L^{\text{MFMC}} \to \frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^i - \beta_{L-1} \frac{1}{m_L} \sum_{i=1}^{m_L} Z_{L-1}^i + \beta_{L-1}\mu_{L-1} = \widehat{\mu}_L^{\text{CV}},$$

where $\widehat{\mu}_L^{\text{CV}}$ is the single CV estimator with control variate Z_{L-1} with m_L samples. □

3.5 Approximate Control Variates

We describe a different control variate framework introduced in [62] in this section. The basic observation of the authors is that the MFMC and the Multilevel Monte Carlo estimator in Section 3.6 do not achieve the same lower variance bound as the control variate estimator. They propose several estimators which do not have this disadvantage.

Definition 3.36 (ACV–IS estimator [62, Definition 2])**.** For $m_1, \dots, m_L \in \mathbb{N}$ and $\beta \in \mathbb{R}^{L-1}$ we define the *Approximate Control Variate Independent Samples (ACV–IS) estimator*

$$\widehat{\mu}_L^{\text{ACV–IS}} := \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^{i,L} - \sum_{\ell=1}^{L-1}\beta_\ell\left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} - \frac{1}{m_\ell}\left(\sum_{i=1}^{m_L} Z_\ell^{i,L} + \sum_{i=m_L+1}^{m_\ell} Z_\ell^{i,\ell}\right)\right). \qquad \diamond$$

We want to choose the coefficients β such that the variance of the estimator is minimized. In contrast to the MFMC estimator, the optimal coefficients β now depend on the number of samples m. We denote the Hadamard or element wise product of two matrices with $\circ$.

Lemma 3.37 (Bias, variance [62, Theorem 3])**.** Define the matrix $F^{IS} \in \mathbb{R}^{(L-1)\times(L-1)}$

$$F^{IS} := (F_{\ell j}^{IS})_{\ell,j=1}^{L-1}, \quad F_{\ell j}^{IS} := \begin{cases} \frac{(m_\ell - m_L)(m_j - m_L)}{m_\ell m_j}, & \text{if } \ell \neq j, \\ \frac{m_\ell - m_L}{m_\ell}, & \text{if } \ell = j. \end{cases}$$

Moreover, let $m_\ell \geq m_L$ for all $\ell \in \{1, \dots, L-1\}$ and for $I := \{1, \dots, L-1\}$ assume that $C_{I,I}$ is positive definite. Then the ACV–IS estimator is an unbiased estimator for μ_L and the variance minimizing choice for β is

$$\beta^* = (F^{IS} \circ C_{I,I})^{-1}(\operatorname{diag}(F^{IS}) \circ C_{I,L}). \tag{3.49}$$

The minimal variance at β^* satisfies

$$\mathbb{V}\left[\widehat{\mu}_L^{\text{ACV–IS}}\right] = \frac{1}{m_L}(\sigma_L^2 - (\operatorname{diag}(F^{IS}) \circ C_{I,L})^T (F^{IS} \circ C_{I,I})^{-1}(\operatorname{diag}(F^{IS}) \circ C_{I,L})). \tag{3.50}$$

Proof. We refer to the proof of [62, Theorem 3]. □

The authors of [62, Definition 3] also define a modification of the MFMC estimator, where the first average of the control variates is now only summed up over the first m_L samples.

Definition 3.38 (ACV–MF estimator [62, Definition 3])**.** For $m_1, \dots, m_L \in \mathbb{N}$ and $\beta \in \mathbb{R}^{L-1}$ we define the *Approximate Control Variate Multifidelity (ACV–MF) estimator*

$$\widehat{\mu}_L^{\text{ACV–MF}} := \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=1}^{L-1}\beta_\ell\left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^i - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^i\right). \qquad \diamond$$

Lemma 3.39 (Bias, variance [62, Theorem 4])**.** Define the matrix $F^{MF} \in \mathbb{R}^{(L-1)\times(L-1)}$

$$F^{MF} := (F_{\ell j}^{MF})_{\ell,j=1}^{L-1}, \quad F_{\ell j}^{MF} := \begin{cases} \frac{\min\{m_\ell, m_j\} - m_L}{\min\{m_\ell, m_j\}}, & \text{if } \ell \neq j, \\ \frac{m_\ell - m_L}{m_\ell}, & \text{if } \ell = j. \end{cases}$$

Moreover, let $m_\ell \geq m_L$ for all $\ell \in \{1, \dots, L-1\}$ and for $I := \{1, \dots, L-1\}$ assume that $C_{I,I}$ is positive definite. Then the ACV–MF estimator is an unbiased estimator for μ_L and the variance minimizing choice for β is

$$\beta^* = (F^{MF} \circ C_{I,I})^{-1}(\operatorname{diag}(F^{MF}) \circ C_{I,L}).$$

The minimal variance at β^* satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–MF}}\big] = \frac{1}{m_L}(\sigma_L^2 - (\operatorname{diag}(F^{MF}) \circ C_{I,L})^T (F^{MF} \circ C_{I,I})^{-1}(\operatorname{diag}(F^{MF}) \circ C_{I,L})).$$

Proof. We refer to the proof of [62, Theorem 4]. □

We continue with the next estimator which is a combination of two control variate estimators.

Definition 3.40 (ACV–KL estimator [62, Definition 4])**.** Let $m_1, \dots, m_L \in \mathbb{N}$ and $\beta \in \mathbb{R}^{L-1}$. Then for $K, N \in \{1, \dots, L\}$ with $N \geq K$ we define the *Approximate control variate KL (ACV–KL) estimator*

$$\begin{aligned}\widehat{\mu}_L^{\text{ACV–KL}} := \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^i &- \sum_{\ell=K}^{L-1} \beta_\ell \left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^i - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^i\right) \\ &- \sum_{\ell=1}^{K-1} \beta_\ell \left(\frac{1}{m_N}\sum_{i=1}^{m_N} Z_\ell^i - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^i\right).\end{aligned} \tag{3.51}$$

◇

The KL in the name ACV–KL stems from the original parameters K, L used in [62]. The idea behind this estimator is to use the ACV–MF estimator with the last control variates $Z_K, \dots, Z_{L-1}$ and then use a CV scheme with the remaining control variates $Z_1, \dots, Z_{K-1}$. These control variates are typically cheaper and this allow us to estimate the expression $\frac{1}{m_N}\sum_{i=1}^{m_N} Z_\ell^i$ with $m_N \geq m_L$ samples, which reduces the variance. In particular, for $N = L$ or $K = 1$ the ACV–KL estimator is equal to the ACV–MF estimator.

Lemma 3.41 (Bias, variance [62, Theorem 6])**.** Define the matrix $F^{KL} := (F_{\ell j}^{KL})_{\ell,j=1}^{L-1} \in \mathbb{R}^{(L-1)\times(L-1)}$ such that

$$F_{\ell j}^{KL} := \begin{cases} \frac{\min\{m_\ell, m_j\} - m_L}{\min\{m_\ell, m_j\}}, & \text{if } \ell, j \geq K, \\ m_L \frac{(m_\ell - m_N)(m_j - m_N) - m_N(\min\{m_\ell, m_j\} - m_N)}{m_j m_\ell m_N}, & \text{if } \ell, j < K, \\ m_L \frac{m_\ell - m_N}{m_\ell m_N}, & \text{if } N > \ell \geq K, j < K, \\ m_L \frac{m_j - m_N}{m_j m_N}, & \text{if } N > j \geq K, \ell < K, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, let $m_\ell \geq m_L$ for all $\ell \in \{1, \dots, L-1\}$, $m_\ell > m_N$ for all $\ell \in \{1, \dots, N-1\}$ and for $I := \{1, \dots, L-1\}$ assume that $C_{I,I}$ is positive definite. Then the ACV–KL estimator is an unbiased estimator for μ_L and the variance minimizing choice for β is

$$\beta^* = (F^{KL} \circ C_{I,I})^{-1}(\operatorname{diag}(F^{KL}) \circ C_{I,L}).$$

The minimal variance at β^* satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–KL}}\big] = \frac{1}{m_L}(\sigma_L^2 - (\operatorname{diag}(F^{KL}) \circ C_{I,L})^T (F^{KL} \circ C_{I,I})^{-1}(\operatorname{diag}(F^{KL}) \circ C_{I,L})).$$

Proof. We refer the reader to the proof of [62, Theorem 6]. □

We give the reader an intuition for the particular expression of β^* and the variance for the ACV–IS estimator. A similar result also holds for the ACV–MF and ACV–KL estimators. We define the *sample modified covariance matrix* as

$$C^m := \frac{1}{m_L}\begin{pmatrix} F^{IS} & \operatorname{diag}(F^{IS}) \\ \operatorname{diag}(F^{IS})^T & 1 \end{pmatrix} \circ C. \tag{3.52}$$

With this notation we rewrite β^* (3.49) and $\mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–IS}}\big]$ in (3.50) to obtain an analogous result to Lemma 3.17 for the multiple CV estimator in terms of the Schur complement

$$\begin{aligned} \beta^* &= (C^m_{I,I})^{-1} C^m_{I,L}, \\ \mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–IS}}\big] &= C^m / C^m_{I,I}. \end{aligned}$$

A possible interpretation, at least in terms of the variance, is that the ACV estimators are CV estimators that use control variates that depend on the number of samples $m_1, \dots, m_L$. For $\ell \in \{1, \dots, L-1\}$ we define

$$\widetilde{Z}_\ell := \frac{1}{m_L^{1/2}} \left[\frac{m_\ell - m_L}{m_\ell} Z_\ell + \left(\frac{m_\ell - m_L}{m_\ell} - \left(\frac{m_\ell - m_L}{m_\ell} \right)^2 \right)^{1/2} \xi_\ell \right], \tag{3.53}$$

where the $\xi_1, \dots, \xi_{L-1}$ are independent copies of $Z_1, \dots, Z_{L-1}$. We further define $\widetilde{Z}_L := Z_L / m_L^{1/2}$. It is now straightforward to verify that the covariance matrix of $\widetilde{Z}_1, \dots, \widetilde{Z}_L$ is equal to (3.52). Furthermore, for $m_\ell > m_L$ this covariance matrix is obviously invertible if C is invertible and thus the expression $(F^{IS} \circ C_{I,I})^{-1}$ in (3.49) is well defined.

Sample allocation. The ACV–IS, ACV–MF and ACV–KL estimators all depend on the number of samples $m_1, \dots, m_L$ and the ACV–KL estimator additionally depends on the parameters K and N. For a positive budget $\mathbb{W}^{\text{budget}} > 0$ the authors of [62] solve the relaxed sample allocation problem

$$\begin{aligned} \min_{m_1,\dots,m_L \in \mathbb{R}} \quad & J(m) := \mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–IS}}\big] \\ \text{such that} \quad & \sum_{\ell=1}^{L} m_\ell w_\ell = \mathbb{W}^{\text{budget}}, \\ & m_\ell \geq m_L \qquad \text{for all } \ell \in \{1, \dots, L-1\}, \\ & m_L \geq 0. \end{aligned}$$

The optimization problem for the ACV–MF estimator is similar, we simply replace the variance $\mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–IS}}\big]$ with the counterpart $\mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–MF}}\big]$. For the ACV–KL estimator we have the integer parameters K and N and thus use categorical optimization. First fix K and N and compute the optimal solution of

$$\begin{aligned} \min_{m_1,\dots,m_L \in \mathbb{R}} \quad & J(m) := \mathbb{V}\big[\widehat{\mu}_L^{\text{ACV–KL}}\big] \\ \text{such that} \quad & \sum_{\ell=1}^{L} m_\ell w_\ell = \mathbb{W}^{\text{budget}}, \\ & m_\ell \geq m_L \qquad \text{for all } \ell \in \{1, \dots, L-1\}, \\ & m_\ell \geq m_N \qquad \text{for all } \ell \in \{1, \dots, N-1\}, \\ & m_L \geq 0. \end{aligned}$$

This is then done for every valid combination of (K, N) and the parameters with the smallest variance is chosen. There is no known analytical form for the optimal sample allocation $m_1, \dots, m_L$ for either of these estimators. Therefore, numerical optimization has to be used.
At this stage we remark that the ACV estimators depend on the ordering of $Z_1, \dots, Z_{L-1}$. In particular, the authors of [62] do not optimize over this ordering and do not perform any model selection. As an example to show why this might be a disadvantage let $Z_1, \dots, Z_{L-1}$ be independent of Z_L. Then the constraint $m_L \leq m_\ell$ requires the evaluation of models that do not reduce the variance but have non–zero costs.

Lower variance bound. The authors of [62] changed a small but important detail for the ACV–IS and the ACV–MF estimator compared to the MFMC estimator. If the number of low fidelity samples goes to infinity, then the contribution of the first term of the control variates $\frac{1}{m_L} \sum_{i=1}^{m_L} Z_\ell^{i,L}$ in Definition 3.36 and $\frac{1}{m_L} \sum_{i=1}^{m_L} Z_\ell^i$ Definition 3.38 for $\ell \in \{1, \dots, L-1\}$ does not go to zero as for MFMC. This allows these estimators to achieve the same lower variance bound as the CV estimators. This was proven in a more general setting by Gorodetsky et al. [62, Theorem 5]. The main idea of the proof is to use the convergence of the matrices $F^{IS} \circ C_{I,I}$ and $F^{MF} \circ C_{I,I}$ to $C_{I,I}$ for an increasing number of low fidelity evaluations. The result is the same lower variance bound as for the CV estimator, which is not a coincidence.

Lemma 3.42 (Almost sure convergence to CV)**.** The ACV–IS and ACV–MF estimator converge almost surely to the multiple CV estimator with control variates $Z_1, \dots, Z_{L-1}$ and m_L samples

$$\lim_{m_1,\dots,m_{L-1}\to+\infty} \widehat{\mu}_L^{\text{ACV–IS}} = \lim_{m_1,\dots,m_{L-1}\to+\infty} \widehat{\mu}_L^{\text{ACV–MF}} = \widehat{\mu}_L^{\text{CV}}[Z_1, \dots, Z_{L-1}]. \tag{3.54}$$

The same statement holds for the ACV–KL estimator with $K = 1$ or $N = L$. For $K > 1$ and $N < L$ the ACV–KL estimator almost surely converges to the multiple CV estimator with control variates $Z_K, \dots, Z_{L-1}$ and m_L samples

$$\lim_{m_1,\dots,m_{L-1}\to+\infty} \widehat{\mu}_L^{\text{ACV–KL}} = \widehat{\mu}_L^{\text{CV}}[Z_K, \dots, Z_{L-1}]. \tag{3.55}$$

Proof. The proof for the ACV–IS estimator is similar to the proof for the ACV–MF estimator, thus we only prove the statement (3.54) for ACV–MF. The estimator reads

$$\widehat{\mu}_L^{\text{ACV–MF}} = \frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=1}^{L-1} \beta_\ell \left(\frac{1}{m_L} \sum_{i=1}^{m_L} Z_\ell^i - \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_\ell^i \right).$$

If we now let $m_\ell \to +\infty$ the terms $\frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_\ell^i \to \mu_\ell$ almost surely, which shows the claim. The claim for the ACV–KL estimator follows since for $K = 1$ or $N = L$ it is equal to the ACV–MF estimator. The proof for $K > 1$ and $N < L$ is straightforward. The terms in the sum $\sum_{\ell=1}^{K-1}$ in (3.51) converge almost surely to zero which shows the claim. □

Another method to derive this result, at least for the ACV–IS estimator, is the interpretation as CV estimator with control variates $\widetilde{Z}_1, \dots, \widetilde{Z}_{L-1}$ defined in (3.53). For $m_L = 1$ and $m_\ell \to +\infty$ for $\ell \in \{1, \dots, L-1\}$ the control variates almost surely converge, that is $\widetilde{Z}_\ell \to Z_\ell$. A straightforward consequence is that the lower variance bound for ACV estimators is equal to the lower variance bound for the CV estimators.

Theorem 3.43 (Lower variance bound)**.** Let $K = 1$ or $N = L$ for the ACV–KL estimator. Then the lower variance bound for ACV is

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\text{ACV–IS}}] = \mathbb{V}^{\min}[\widehat{\mu}_L^{\text{ACV–MF}}] = \mathbb{V}^{\min}[\widehat{\mu}_L^{\text{ACV–KL}}] = \mathbb{V}^{\min}_{\{1,\dots,L-1\}}. \tag{3.56}$$

For $K > 1$ and $N < L$ the lower variance bound for ACV–KL is

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\text{ACV–KL}}] = \mathbb{V}^{\min}_{\{K,\dots,L-1\}}. \qquad \diamond$$

3.6 Multilevel Monte Carlo

The Multilevel Monte Carlo (MLMC) estimator has recently become popular due to its asymptotic complexity improvement over the MC estimator. Giles [56] introduced this estimator for computing the expectation of the solution of a stochastic differential equation. The authors of [31] extended this result to a PDE with a random diffusion coefficient. A summary of the method and extensions are listed in [57]. Let us denote $Z_0 := 0$, $w_0 := 0$ and $\mu_0 := 0$.

Definition 3.44 (MLMC estimator [56, Section 2])**.** For $m_1, \dots, m_L$ we define *MLMC estimator*

$$\widehat{\mu}_L^{\text{MLMC}} := \sum_{\ell=1}^{L} \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} (Z_\ell^{i,\ell} - Z_{\ell-1}^{i,\ell}). \tag{3.57}$$

$\diamond$

The estimator differs from a CV estimator in the sense that there are no coefficients β to choose. Furthermore, the samples for every difference in (3.57) are independent. In particular, m_ℓ is **not** the number of evaluations of Z_ℓ but rather of $Z_\ell - Z_{\ell-1}$. The total number of evaluations of Z_ℓ is $m_\ell + m_{\ell+1}$ for $\ell \in \{1, \dots, L-1\}$ and of Z_L is m_L.
It is not immediately obvious why the MLMC estimator is constructed as a telescoping sum, however the reasons will become clear when we examine its asymptotic properties. Let us write down the bias and variance of this estimator.

Lemma 3.45 (Bias, variance [56, Section 2])**.** The MLMC estimator is an unbiased estimator for μ_L with variance

$$\mathbb{V}\big[\widehat{\mu}_L^{\text{MLMC}}\big] = \sum_{\ell=1}^{L} \frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{m_\ell}. \tag{3.58}$$

Proof. We use the properties of the telescoping sum, the linearity of the expectation and $\mu_0 = 0$ to conclude the unbiasedness

$$\widehat{\mu}_L^{\text{MLMC}} = \sum_{\ell=1}^{L} \mathbb{E}[Z_\ell - Z_{\ell-1}] = \mu_L - \mu_0 = \mu_L.$$

We use the independence of the random variables w.r.t. the levels ℓ and the sample index i together with the standard MC variance estimate to conclude (3.58)

$$\mathbb{V}\big[\widehat{\mu}_L^{\text{MLMC}}\big] = \sum_{\ell=1}^{L} \mathbb{V}\left[\frac{1}{m_\ell} \sum_{i=1}^{m_\ell} (Z_\ell^{i,\ell} - Z_{\ell-1}^{i,\ell})\right] = \sum_{\ell=1}^{L} \frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{m_\ell}. \qquad \square$$

Sample allocation. Let us continue with the sample allocation for MLMC. The goal is to find the number of samples $m_1, \dots, m_L$ that minimize the variance given a fixed computational budget $\mathbb{W}^{\text{budget}} > 0$. The relaxed sample allocation problem where we allow fractional samples reads

$$\min_{m_1,\dots,m_L \in \mathbb{R}} \quad J(m) := \mathbb{V}\big[\widehat{\mu}_L^{\text{MLMC}}\big] = \sum_{\ell=1}^{L} \frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{m_\ell} \tag{3.59}$$

$$\text{such that} \quad \sum_{\ell=1}^{L} m_\ell (w_\ell + w_{\ell-1}) = \mathbb{W}^{\text{budget}}, \tag{3.60}$$

$$m_\ell \geq 0 \qquad \text{for all } \ell \in \{1, \dots, L\}. \tag{3.61}$$

This minimization problem has a unique solution which can be computed analytically. The solution was obtained by Giles [56] and generalized by Cliffe et al. [31] to accommodate the PDE setting.

Theorem 3.46 (Optimal sample allocation [31, Section 2.1])**.** Let $\mathbb{V}[Z_\ell - Z_{\ell-1}] > 0$ for all $\ell \in \{1, \dots, L\}$. Then the optimal sample allocation m^* of (3.59)-(3.61) is

$$m_\ell^* = \frac{\mathbb{W}^{\text{budget}}}{\sum_{\ell=1}^{L} \left(\mathbb{V}[Z_\ell - Z_{\ell-1}](w_\ell + w_{\ell-1})\right)^{1/2}} \left(\frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{w_\ell + w_{\ell-1}}\right)^{1/2}.$$

The variance at this minimizer satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\text{MLMC}}\big] = \frac{1}{\mathbb{W}^{\text{budget}}} \left(\sum_{\ell=1}^{L} \left(\mathbb{V}[Z_\ell - Z_{\ell-1}](w_\ell + w_{\ell-1})\right)^{1/2}\right)^2. \tag{3.62}$$

Proof. The proof is standard convex analysis [20, Section 5]. We however, want to repeat the main ideas of it since a proof with similar structure will appear in a more complicated setting. First, we argue that a minimizer exists. We use $\mathbb{V}[Z_\ell - Z_{\ell-1}] > 0$ and $m_\ell \geq 0$ to show

$$\lim_{m_\ell \to 0} J(m) = +\infty.$$

Since $J(m)$ is continuous for $m_\ell > 0$, we conclude that $m_\ell \geq c > 0$ for all $\ell \in \{1, \dots, L\}$ for some small c. Furthermore, the cost constraint (3.60) and $w_\ell > 0$ ensures that

$$m_\ell (w_\ell + w_{\ell-1}) \leq \mathbb{W}^{\text{budget}}$$

and thus $m_\ell \leq c$ for all $\ell \in \{1, \dots, L\}$ for sufficiently large c. The continuous function J attains its minimum over a compact set, thus m^* exists. The uniqueness follows since J is strictly convex restricted to linear cost constraint. This can be seen by looking at the Hessian of J, which is a diagonal matrix that is strictly positive definite

$$H_J(m) = (H_J(m)_{\ell j})_{\ell,j=1}^{L}, \quad H_J(m)_{\ell j} = \begin{cases} 2\mathbb{V}[Z_\ell - Z_{\ell-1}] m_\ell^{-3}, & \text{if } \ell = j, \\ 0, & \text{if } \ell \neq j. \end{cases}$$

This also shows that (3.59)-(3.61) is a convex optimization problem, since the constraints are all affine. The *Karush-Kuhn-Tucker* (KKT) conditions are thus sufficient for a minimizer. There exists the Lagrange multiplier $\lambda^{\mathbb{W}} \in \mathbb{R}$ associated with the cost constraint

and $\lambda_\ell \in \mathbb{R}$ associated with the positivity constraint of m_ℓ for every $\ell \in \{1, \dots, L\}$ such that the KKT conditions are

$$\begin{aligned} -\frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{m_\ell^2} + \lambda^{\mathbb{W}}(w_\ell + w_{\ell-1}) - \lambda_\ell &= 0 && \text{for all } \ell \in \{1, \dots, L\}, \\ \sum_{\ell=1}^{L} m_\ell(w_\ell + w_{\ell-1}) &= \mathbb{W}^{\text{budget}}, \\ m_\ell &\geq 0 && \text{for all } \ell \in \{1, \dots, L\}, \\ \lambda_\ell m_\ell &= 0 && \text{for all } \ell \in \{1, \dots, L\}, \\ \lambda_\ell &\geq 0 && \text{for all } \ell \in \{1, \dots, L\}. \end{aligned} \tag{3.63}$$

Observe that we already concluded $m_\ell > 0$ and thus by the complementary slackness condition $\lambda_\ell = 0$. Therefore, we have to solve the system

$$\begin{aligned} -\frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{m_\ell^2} + \lambda^{\mathbb{W}}(w_\ell + w_{\ell-1}) &= 0 && \text{for all } \ell \in \{1, \dots, L\}, \\ \sum_{\ell=1}^{L} m_\ell(w_\ell + w_{\ell-1}) &= \mathbb{W}^{\text{budget}}. \end{aligned}$$

We further conclude that $\lambda^{\mathbb{W}} > 0$ and thus for all $\ell \in \{1, \dots, L\}$

$$m_\ell = \left(\frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{\lambda^{\mathbb{W}}(w_\ell + w_{\ell-1})} \right)^{1/2}.$$

The cost constraint (3.60) now shows

$$\lambda^{\mathbb{W}} = \frac{1}{(\mathbb{W}^{\text{budget}})^2} \left(\sum_{\ell=1}^{L} (\mathbb{V}[Z_\ell - Z_{\ell-1}](w_\ell + w_{\ell-1}))^{1/2} \right)^2.$$

This then yields the optimal value for m_ℓ^* for all $\ell \in \{1, \dots, L\}$

$$m_\ell^* = \frac{\mathbb{W}^{\text{budget}}}{\sum_{\ell=1}^{L} (\mathbb{V}[Z_\ell - Z_{\ell-1}](w_\ell + w_{\ell-1}))^{1/2}} \left(\frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{w_\ell + w_{\ell-1}} \right)^{1/2}.$$

The variance at the minimizer m^* satisfies

$$J(m^*) = \frac{1}{\mathbb{W}^{\text{budget}}} \left(\sum_{\ell=1}^{L} (\mathbb{V}[Z_\ell - Z_{\ell-1}](w_\ell + w_{\ell-1}))^{1/2} \right)^2 = \lambda^{\mathbb{W}} \mathbb{W}^{\text{budget}}. \qquad \square$$

It is straightforward to generalize the theorem if $\mathbb{V}[Z_\ell - Z_{\ell-1}] = 0$ for some $\ell \in \{1, \dots, L\}$. We remove the model ℓ from the estimator without affecting the accuracy, which corresponds to $m_\ell^* = 0$. The result (3.62) shows the inversely proportional relationship between $\mathbb{V}\left[\widehat{\mu}_L^{\text{MLMC}}\right]$ and $\mathbb{W}^{\text{budget}}$, similar to the result for MFMC in Theorem 3.26.

Graph based model selection. The variance of the MLMC estimator (3.62) shows that the cost is not invariant under reordering of the models. Let us formulate a model selection problem similar to the one for the MFMC estimator in Definition 3.28.

Definition 3.47 (Model selection problem)**.** The model selection for MLMC is to minimize

$$\min_P \quad J(P) := \sum_{\ell=1}^{|P|} \big(\mathbb{V}\big[Z_{P_\ell} - Z_{P_{\ell-1}}\big](w_{P_\ell} + w_{P_{\ell-1}})\big)^{1/2}, \tag{3.64}$$

where $P = (P_1, P_2, \dots, P_{|P|-1}, L)^T$ is a vector of length $|P| \leq L$ with $P_\ell \in \{1, \dots, L\}$ and we defined $w_{P_0} := 0$. ◇

The cost function J in (3.64) is up to constant equal to the square root of the variance of the MLMC estimator in (3.62) for $P = (1, \dots, L)^T$. Definition 3.47 allows us to reorder the models and if necessary, not use some of them.
The goal is to find the optimal model selection P that minimizes the variance and we show that this problem is equivalent to a shortest path problem. Let us define the directed, edge weighted graph $G := (V, E, Q)$ defined as

$$\begin{aligned} V &:= \{0, \dots, L\}, \\ E &:= V \times V, \\ Q &:= (q_{ij})_{i,j=0}^{L}, \quad q_{ij} := (\mathbb{V}[Z_j - Z_i](w_j + w_i))^{1/2}. \end{aligned} \tag{3.65}$$

A shortest path from 0 to L now describes the best model selection and can be computed efficiently, i.e. by using Dijkstra's algorithm, see [13, Section 1.5] or [2, Section 7.4].

Lemma 3.48 (Shortest path model selection)**.** Finding a minimizer of (3.64) is equivalent to finding a shortest path from 0 to L in G defined in (3.65).

Proof. The model selection $P = (P_1, P_2, \dots, P_{|P|-1}, L)^T$ is a valid path from 0 to L in G if we add zero in front $P' = (0, P)^T$. Similarly, every path from 0 to L in G is a valid model selection if we drop the first entry. All that remains is to show that the cost of any path P from 0 to L is equal to $J(P)$ in (3.64). This however, follows from the definition of the edge weights q_{ij} in (3.65). □

Asymptotic complexity. The asymptotic cost for the MLMC estimator is an improvement over the MC estimator and is equal to the cost of the MFMC estimator in Theorem 3.31. The proof is also quite similar. In particular, the MLMC estimator achieves the optimal $\mathcal{O}(\varepsilon^{-2})$ complexity if the variance reduction rate γ_{Var} is larger than the cost increase rate γ_{Cost}.

Theorem 3.49 (Asymptotic cost [56, Theorem 3.1], [31, Theorem 1])**.** Let the following assumptions be true for all $\ell \in \mathbb{N}$

$$\text{Bias}(Z_\ell) \leq c 2^{-\gamma_{\text{Bias}}\ell}, \tag{3.66}$$

$$\mathbb{V}[Z_\ell - Z_{\ell-1}] \leq c 2^{-\gamma_{\text{Var}}\ell}, \tag{3.67}$$

$$w_\ell \leq c 2^{\gamma_{\text{Cost}}\ell}. \tag{3.68}$$

Then for all $\varepsilon \in (0, 1/e]$ there exist a final level L and a number samples $m_1, \dots, m_L$ such that the cost to achieve $\text{MSE}(\widehat{\mu}_L^{\text{MLMC}}) \leq \varepsilon^2$ is bounded

$$\mathbb{W}\big[\widehat{\mu}_L^{\text{MLMC}}\big] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} + c\begin{cases} \varepsilon^{-2}, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}}. \end{cases}$$

Proof. The proof is analogous to the proof for the MFMC estimator in Theorem 3.31. We only have to verify that the variance satisfies the expression

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MLMC}}\big] \leq \frac{c}{\mathbb{W}^{\mathrm{budget}}} \left(\sum_{\ell=1}^{L} 2^{(\gamma_{\mathrm{Cost}}-\gamma_{\mathrm{Var}})\ell/2} \right)^2,$$

which directly follows if we combine (3.62) with (3.67) and (3.68). □

Lower variance bound. The lower variance bound of MLMC does not reach the bound $\mathbb{V}^{\min}_{\{L-1\}}$ as for a single CV estimator, since there is no coefficient β to choose. This result can be found in [62, Section 2.4], we however provide a different proof.

Theorem 3.50 (Lower variance bound)**.** The MLMC estimator almost surely converges

$$\lim_{m_1,\dots,m_{L-1}\to+\infty} \widehat{\mu}_L^{\mathrm{MLMC}} = \frac{1}{m_L} \sum_{i=1}^{m_L} (Z_L^i - Z_{L-1}^i) + \mu_{L-1}. \tag{3.69}$$

In particular, the lower variance bound for the MLMC estimator satisfies

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\mathrm{MLMC}}] = \mathbb{V}[Z_L - Z_{L-1}] \geq \mathbb{V}^{\min}_{\{L-1\}}. \tag{3.70}$$

Proof. All MC estimators in (3.57) converge almost surely to their mean

$$\lim_{m_1,\dots,m_{L-1}\to+\infty} \widehat{\mu}_L^{\mathrm{MLMC}} = \frac{1}{m_L} \sum_{i=1}^{m_L} (Z_L^i - Z_{L-1}^i) + \sum_{\ell=1}^{L-1} (\mu_\ell - \mu_{\ell-1}),$$

which shows (3.69) since this is a telescoping sum with $\mu_0 = 0$. The bound (3.70) now follows since the right estimator in (3.69) is a single CV estimator with control variate Z_{L-1}, m_L samples and potentially suboptimal coefficient $\beta = 1$. □

Let us briefly outline a derivation of the expression (3.70) by decreasing the cost of the low fidelity models to zero. We use (3.62) with $\mathbb{W}^{\mathrm{budget}} = w_L$ to obtain

$$\begin{aligned} \lim_{w_1,\dots,w_{L-1}\to 0} \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MLMC}}\big] &= \lim_{w_1,\dots,w_{L-1}\to 0} \frac{1}{\mathbb{W}^{\mathrm{budget}}} \left(\sum_{\ell=1}^{L} (\mathbb{V}[Z_\ell - Z_{\ell-1}](w_\ell + w_{\ell-1}))^{1/2} \right)^2 \\ &= \mathbb{V}[Z_L - Z_{L-1}]. \end{aligned}$$

3.7 Other multilevel methods

Multi-index Monte Carlo. The Multi-index Monte Carlo estimator [68] is an extension of the MLMC estimator such that the telescoping sum idea is used in multiple directions. Let us assume that we have the models $Z_{\ell,j}$ for $\ell, j \in \{1, \dots, L\}$ available. Then define the two difference operators in the first and the second direction

$$\Delta_1 Z_{i,j} := \begin{cases} Z_{1,j}, & \text{if } i = 1, \\ Z_{i,j} - Z_{i-1,j}, & \text{if } i \neq 1, \end{cases} \quad \Delta_2 Z_{i,j} := \begin{cases} Z_{i,1}, & \text{if } j = 1, \\ Z_{i,j} - Z_{i,j-1}, & \text{if } j \neq 1. \end{cases}$$

The estimator for $\mathbb{E}[Z_{L,L}]$ now sums up over all low fidelity models

$$\widehat{\mu}_{L,L} := \sum_{\ell,j=1}^{L} \frac{1}{m_{\ell,j}} \sum_{i=1}^{m_{\ell,j}} \Delta_2 \Delta_1 Z_{\ell,j}^{i,j,\ell}.$$

This is an unbiased estimator for $\mathbb{E}[Z_{L,L}]$ with variance

$$\mathbb{V}\big[\widehat{\mu}_{L,L}\big] = \sum_{\ell,j=1}^{L} \frac{\mathbb{V}[\Delta_2 \Delta_1 Z_{\ell,j}]}{m_{\ell,j}}.$$

This expression is equivalent to the expression for the MLMC estimator (3.58) except that we sum over two dimensions. Importantly, this method is viable if the variance term exhibits mixed regularity in the following sense

$$\mathbb{V}[\Delta_2 \Delta_1 Z_{\ell,j}] = c 2^{-\gamma_{\mathrm{Var}2}\ell} 2^{-\gamma_{\mathrm{Var}1} j},$$

where $\gamma_{\mathrm{Var}2}, \gamma_{\mathrm{Var}1}$ are positive rates w.r.t. the respective dimensions.
It is possible and advisable to not sum up over all $\ell, j \in \{1, \dots, L\}$ and instead carefully select the indices. The authors show in [68, Remark 2.2] that with an improved selection and under some regularity assumptions, the convergence rate is, up to logarithmic factor, independent of the dimension d of D for the PDE setting in Example 3.12. These results however, require mixed regularity assumptions w.r.t. the bias, variance and cost.

Adaptive methods. Adaptive methods seek to optimize the low fidelity models $Z_1, \dots, Z_{L-1}$ in a certain way to reduce the variance. The method of Peherstorfer [105] uses a Gaussian process to generate a cheap low fidelity model and uses this model as control variate to estimate the expectation of the high fidelity model. The tradeoff here is that the low fidelity model may be inaccurate if it is cheap.
Other adaptive methods do not consider a fixed hierarchy $Z_1, \dots, Z_{L-1}$ and use a refinement scheme that depends on the particular sample, see [44, 74, 80]. This exploits features of the underlying QoI that may be sample dependent such as a localized forcing term, where the location is random and differs significantly for different realizations. Adaptive mesh refinement is then used to obtain sample dependent approximations of Z.

Chapter 4

Multilevel best linear unbiased estimators

In this chapter we reformulate sampling based estimation as linear regression problem and generalized linear least squares problem. These problems and their respective solutions are well known in the mathematical literature [8, 64, 69, 96, 114, 116, 142]. Therefore, a lot of work in this chapter consists of translating sampling based estimation to the language of least squares problems. We restrict ourselves to a solution in the class of linear unbiased estimators, where we form a linear combination of the samples. This class includes the estimators presented in Chapter 3. The Gauss–Markov–Aitken Theorem asserts the existence and uniqueness of the best linear unbiased estimator (BLUE), where "best" means variance minimal. In this context the term "multilevel" means that we use the samples of the low fidelity models $Z_1, \dots, Z_{L-1}$ to estimate the expectation of the high fidelity model Z_L. The contents of this chapter is organized as follows:

- We derive the BLUE from two different perspectives in Section 4.1. First, we proceed similar to the CV estimator by defining a general linear unbiased estimator. Then we choose the coefficients to satisfy a bias constraint and to be variance minimal, which in turn leads to the BLUE. The second approach reformulates the estimation problem as linear regression problem, where we apply the Gauss–Markov–Aitken Theorem to obtain the BLUE. Both derivations lead to the same result, however the latter one is more insightful.

- We continue with the lower variance bound for the BLUE in Section 4.2. Since the BLUE is the linear unbiased estimator with the smallest variance, this bound is a lower bound for every linear unbiased estimator. As it turns out the bound is sharp and equal to the bound for the CV and ACV estimators.

- Linear unbiased estimators are formed by linear combinations of the samples. In Section 4.3 we view the samples as basis functions of a suitable subspace V. The BLUE consists of the best approximation in V of an element in a larger space which we obtain from the bias constraint plus a residual orthogonal to V. This viewpoint emphasizes the samples and the QoIs $Z_1, \dots, Z_L$ and not the coefficients of the linear combination.

- We view the estimators of Chapter 3 as linear unbiased estimator in Section 4.4 and show that some of them are the BLUE under certain circumstances. We further define the full coupling estimators, which are the BLUEs for a special sample allocation.

- Some results of this chapter are verified numerically in Section 4.5, where we have samples of a QoI which is a monomial or a monomial plus noise. The goal is to combine these samples in a linear fashion to reduce the variance of the estimate for the mean of the high fidelity model μ_L. We further verify that $\mathbb{V}^{\min}$ in Definition 3.19 is the lower variance bound.

Throughout this chapter we estimate a linear combination of the means of the discretized QoIs $\mathbb{E}[Z_1], \dots, \mathbb{E}[Z_L]$ and not of the actual mean $\mathbb{E}[Z]$ we are interested in. This is a

restriction that makes sense for the purpose of this chapter and we defer the estimation of the exact mean $\mathbb{E}[Z]$ to Chapter 6. This chapter contains results from [126].

4.1 Estimation as linear regression

Linear unbiased estimators. We estimate the mean $\mu = (\mathbb{E}[Z_1], \dots, \mathbb{E}[Z_L])^T$ by linearly combining samples of $Z_1, \dots, Z_L$. We define subsets of these QoIs as model groups, where we later independently sample from them. We further define a restriction and prolongation operator.

Definition 4.1 (Model group, restriction, prolongation)**.** For $K := 2^L - 1$ let $S^1, \dots, S^K$ be the K non–empty, disjoint subsets of $\{1, \dots, L\}$. We call each S^k a *model group*. For each model group the *restriction* R^k is defined such that for all $v \in \mathbb{R}^L$

$$R^k v := v_{S^k} \in \mathbb{R}^{|S^k|}.$$

We define the *prolongation* as transpose of the restriction $P^k := (R^k)^T$. ◇

We simplify the notation by deliberately not specifying which subset the model group S^k for a specific k is unless it is contextually required. We now define estimators that linearly combine samples with coefficients that are deterministic and thus implicitly define the bias parameter. For every $k \in \{1, \dots, K\}$ we have m_k i.i.d. samples of the QoIs with indices in S^k and denote them as usual with $Z^{1,k}_{S^k}, \dots, Z^{m_k,k}_{S^k}$.

Definition 4.2 (Linear unbiased estimator)**.** For deterministic coefficients $\beta^{i,k} \in \mathbb{R}^{|S^k|}$ and samples $Z^{i,k}_{S^k}$ we call $\widehat{\mu}_\alpha$ a *linear estimator*

$$\widehat{\mu}_\alpha := \sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T Z^{i,k}_{S^k}.$$

The *bias (parameter)* α is defined in terms of the coefficients

$$\alpha := \sum_{k=1}^{K} P^k \sum_{i=1}^{m_k} \beta^{i,k}. \tag{4.1}$$

The estimator $\widehat{\mu}_\alpha$ is a *linear unbiased estimator* (for $\alpha^T \mu$). For unit vectors $e_\ell \in \mathbb{R}^L$ we abbreviate $\widehat{\mu}_\ell := \widehat{\mu}_{e_\ell}$. ◇

We call α the bias since it determines the bias for estimating zero. The definition (4.1) ensures that $\widehat{\mu}_\alpha$ is an unbiased estimator for $\alpha^T \mu$ irrespective of the actual value of μ, which is required since we do not know any entry of the mean μ. This becomes clear if we equivalently reformulate (4.1)

$$\mathbb{E}[\widehat{\mu}_\alpha] = \sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T \mathbb{E}\left[Z^{i,k}_{S^k}\right] = \sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T R^k \mu = \alpha^T \mu \qquad \text{for all } \mu \in \mathbb{R}^L.$$

It is often helpful to specifically select the coefficient in front of the ℓ-th level

$$\beta^{i,k}_{[\ell]} := e_\ell^T P^k \beta^{i,k}, \tag{4.2}$$

where the coefficient is zero $\beta_{[\ell]}^{i,k} = 0$ if the model is not in the group $\ell \notin S^k$. Notice that in general $\beta_{[\ell]}^{i,k} \neq \beta_\ell^{i,k}$ and that $\beta_\ell^{i,k}$ is ill–formed if $\ell > |S^k|$. For the set $I \subseteq \{1, \ldots, L\}$ we define

$$\beta_{[I]}^{i,k} := (\beta_{[\ell]}^{i,k})_{\ell \in I} \in \mathbb{R}^{|I|}. \tag{4.3}$$

We allow the number of samples for a model group to be zero, hence we distinguish models that we are required to evaluate, used models and used model groups.

Definition 4.3 (Required models, used models, used model groups)**.** We define *required models* that we have to evaluate

$$U_\alpha := \{\ell \in \{1, \ldots, L\} \,|\, \alpha_\ell \neq 0\}. \tag{4.4}$$

The models that are evaluated at least once are the *used models*

$$U_Z := \{\ell \in \{1, \ldots, L\} \,|\, \text{ there exists } k \in \{1, \ldots, K\} \text{ such that } \ell \in S^k \text{ and } m_k > 0\}. \tag{4.5}$$

The *used model groups* are evaluated at least once

$$U_S := \{k \in \{1, \ldots, K\} \,|\, m_k > 0\}. \tag{4.6}$$

◇

A step in the construction is to choose the coefficients β given a desired bias α and the number of samples. For every ℓ with $\alpha_\ell \neq 0$ we require to have at least one sample of Z_ℓ, since otherwise the estimator is biased for $\mu_\ell \neq 0$.

Lemma 4.4 (Existence of linear unbiased estimators)**.** For a fixed $\alpha \in \mathbb{R}^L$ there exists a linear unbiased estimator $\widehat{\mu}_\alpha$ if and only if $U_\alpha \subseteq U_Z$.

Proof. "⇒" We fix $\ell \in U_\alpha$ and combine this with (4.1)

$$0 \neq \alpha_\ell = e_\ell^T \alpha = \sum_{k=1}^K e_\ell^T P^k \sum_{i=1}^{m_k} \beta^{i,k} = \sum_{k=1}^K \sum_{i=1}^{m_k} \beta_{[\ell]}^{i,k}.$$

Hence there exists $k \in \{1, \ldots, K\}$ such that $\ell \in S^k$ and $m_k > 0$, which shows $\ell \in U_Z$.
"⇐" We verify that (4.1) is satisfied if we average the coefficients as follows

$$\beta_{[\ell]}^{i,k} := \begin{cases} \frac{1_{S^k}(\ell)}{\sum_{k=1}^K m_k 1_{S^k}(\ell)} \alpha_\ell, & \text{if } \ell \in U_Z, \\ 0, & \text{if } \ell \notin U_Z. \end{cases}$$

For $\ell \in U_Z$ the denominator in this definition is positive and we obtain the desired bias

$$e_\ell^T \sum_{k=1}^K P^k \sum_{i=1}^{m_k} \beta^{i,k} = \sum_{k=1}^K \sum_{i=1}^{m_k} \beta_{[\ell]}^{i,k} = \alpha_\ell \sum_{k=1}^K \sum_{i=1}^{m_k} \frac{1_{S^k}(\ell)}{\sum_{k=1}^K m_k 1_{S^k}(\ell)} = \alpha_\ell.$$

For $\ell \notin U_Z$ we use the assumption $U_\alpha \subseteq U_Z$ and hence $\alpha_\ell = 0$, which is equal to the linear combination of the coefficients since they all satisfy $\beta_{[\ell]}^{i,k} = 0$. □

It is straightforward to extend the notion of a linear unbiased estimator for the entire vector μ or $A\mu$ for any compatible matrix A. This is achieved by separately looking at the rows of A. We abbreviate such an estimator with $\widehat{\mu}_A$ or $\widehat{\mu}$ in the case of $A = I$ the identity matrix. Clearly, a linear unbiased estimator for μ requires us to evaluate all models at least once.

Corollary 4.5 (Existence of linear unbiased estimators)**.** Define the set of required models

$$U_A := \{\ell \in \{1, \dots, L\} \mid \text{ there exists } i \text{ such that } A_{i\ell} \neq 0\}. \qquad \diamond$$

For fixed $A \in \mathbb{R}^{N \times L}$ there exists a linear unbiased estimator $\widehat{\mu}_A$ if and only if $U_A \subseteq U_Z$. In particular, for $A = I \in \mathbb{R}^{L \times L}$ we require $U_Z = \{1, \dots, L\}$. In the next sections and throughout the rest of this thesis we tacitly assume that we have samples such that $U_A \subseteq U_Z$ is satisfied and thus linear unbiased estimators exist.

Best linear unbiased estimator. We now write down the variance of a linear unbiased estimator. The result contains expressions with the covariance matrix of a model group

$$C^k := R^k C P^k \in \mathbb{R}^{|S^k| \times |S^k|}.$$

We assume that the samples are uncorrelated across model groups and that all samples of every model group are also uncorrelated. The random variables $Z_{S^k}^{i,k}$ have the same variance for different sample index i and thus the same information regarding μ_{S^k}. Hence it makes sense to all weigh them exactly the same way $\beta^{i,k} = \beta^{j,k}$ for all $i, j \in \{1, \dots, m_k\}$. We formally derive this.

Lemma 4.6 (Variance, uniform coefficients)**.** The linear unbiased estimator $\widehat{\mu}_\alpha$ has the variance

$$\mathbb{V}[\widehat{\mu}_\alpha] = \sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T C^k \beta^{i,k}. \tag{4.7}$$

For $\beta^k := \sum_{i=1}^{m_k} \beta^{i,k}$ the modified estimator

$$\widehat{\mu}'_\alpha := \sum_{k \in U_S} (\beta^k)^T \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right)$$

uses the same samples, has the same bias and variance not larger than $\widehat{\mu}_\alpha$, that is

$$\mathbb{V}[\widehat{\mu}'_\alpha] = \sum_{k \in U_S} \frac{(\beta^k)^T C^k \beta^k}{m_k} \leq \mathbb{V}[\widehat{\mu}_\alpha]. \tag{4.8}$$

Proof. The variance (4.7) follows from the independence structure of the samples

$$\mathbb{V}[\widehat{\mu}_\alpha] = \mathbb{V}\left[\sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T Z_{S^k}^{i,k} \right] = \sum_{k=1}^{K} \sum_{i=1}^{m_k} \mathbb{V}\left[(\beta^{i,k})^T Z_{S^k}^{i,k} \right] = \sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T C^k \beta^{i,k}.$$

The estimator $\widehat{\mu}'_\alpha$ clearly uses the same samples as $\widehat{\mu}_\alpha$ and is a linear unbiased estimator for $\alpha^T \mu$

$$\begin{aligned} \mathbb{E}[\widehat{\mu}'_\alpha] &= \mathbb{E}\left[\sum_{k=1}^{K} (\beta^k)^T \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right) \right] = \sum_{k=1}^{K} (\beta^k)^T \mu_{S^k} = \sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T \mu_{S^k} \\ &= \mathbb{E}\left[\sum_{k=1}^{K} \sum_{i=1}^{m_k} (\beta^{i,k})^T Z_{S^k}^{i,k} \right] = \mathbb{E}[\widehat{\mu}_\alpha] = \alpha^T \mu. \end{aligned} \tag{4.9}$$

For the variance (4.8) we separately minimize the inner sums of the variance in (4.7) such that we do not change the bias

$$\begin{aligned} \min_{v^1,\dots,v^{m_k}\in\mathbb{R}^{|S^k|}} \quad & J(v^1,\dots,v^{m_k}) := \sum_{i=1}^{m_k} (v^i)^T C^k v^i \\ \text{such that} \quad & \sum_{i=1}^{m_k} v^i = \sum_{i=1}^{m_k} \beta^{i,k} = \beta^k. \end{aligned} \tag{4.10}$$

Clearly, if $m_k = 0$ then the optimization problem is trivial. For $k \in U_S$ and thus $m_k > 0$ this is a convex optimization since C^k is positive semi–definite and the constraint is affine. The KKT conditions are thus necessary and sufficient [20, Section 5]. Let $\lambda \in \mathbb{R}^{|S^k|}$ be the Lagrange multiplier such that the optimal choice for $v^1,\dots,v^{m_k}$ satisfies the KKT conditions

$$\begin{aligned} C^k v^i + \lambda &= 0 && \text{for all } i \in \{1,\dots,m_k\}, \\ \beta^k &= \sum_{i=1}^{m_k} v^i. \end{aligned}$$

We write down a solution of this system

$$\begin{aligned} v^i &= \frac{\beta^k}{m_k} && \text{for all } i \in \{1,\dots,m_k\}, \\ \lambda &= -\frac{C^k}{m_k}\beta^k. \end{aligned} \tag{4.11}$$

We conclude that using uniform coefficients across samples of a model group never increases the variance and thus (4.8) holds. □

The proof of Lemma 4.6 further shows that if the covariance matrix C^k is positive definite, then the optimal choice for $v^1,\dots,v^{m_k}$ in (4.10) is unique with uniform coefficients for every model group.
Notice that we have to make the distinction between $m_k = 0$ and $m_k > 0$, since we otherwise divide by zero in (4.11). The estimator with equal coefficients in front of every sample for each model group is a weighted sum of correlated MC estimators

$$\widehat{\mu}_\alpha := \sum_{k\in U_S} (\beta^k)^T \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right) = \sum_{k\in U_S} \sum_{\ell\in S^k} \beta_{[\ell]}^k \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_\ell^{i,k} \right). \tag{4.12}$$

The inner sum contains correlated MC estimators when changing ℓ and the outer sum contains uncorrelated MC estimators when changing k. The variance of this estimator is

$$\mathbb{V}[\widehat{\mu}_\alpha] = \sum_{k\in U_S} \frac{(\beta^k)^T C^k \beta^k}{m_k} \tag{4.13}$$

and the bias satisfies the simplified expression

$$\alpha = \sum_{k\in U_S} P^k \beta^k. \tag{4.14}$$

The goal is now to choose the coefficients β such that the variance (4.13) is minimized and the bias constraint (4.14) is satisfied. This is a quadratic optimization problem with

positive semi–definite matrices C^k and a linear equality constraint. We define a matrix to express the solution

$$\Psi := \sum_{k \in U_S} m_k P^k (C^k)^{-1} R^k \in \mathbb{R}^{L \times L}$$

and write down its properties.

Lemma 4.7 (Properties of Ψ)**.** The matrix Ψ is well defined if the model group covariance matrices C^k are positive definite. Then Ψ is symmetric and positive semi–definite. Furthermore, Ψ is positive definite if and only if we evaluate every model at least once, that is $U_Z = \{1, \dots, L\}$.

Proof. The well–definedness, symmetry and positive semi–definiteness are easy to verify. Let us now verify the statement for the positive definiteness.
"$\Leftarrow$": Let $v \in \mathbb{R}^L$ be arbitrary such that

$$0 = v^T \Psi v = \sum_{k \in U_S} m_k v^T P^k (C^k)^{-1} R^k v = \sum_{k \in U_S} m_k v_{S^k}^T (C^k)^{-1} v_{S^k}.$$

For $k \in U_S$ we use $m_k > 0$ and that $(C^k)^{-1}$ is positive definite to show

$$v_{S^k} = (v_\ell)_{\ell \in S^k} = 0.$$

For $\ell \in \cup_{k \in U_S} S^k = U_Z$ we conclude $v_\ell = 0$. Since we use all models $U_Z = \{1, \dots, L\}$ we obtain $v = 0$ and thus Ψ is positive definite.
"$\Rightarrow$": Let us assume that the model $\ell \notin U_Z$ is not used, which implies that $R^k e_\ell = 0$ for all $k \in U_S$. Then

$$e_\ell^T \Psi e_\ell = \sum_{k \in U_S} m_k e_\ell^T P^k (C^k)^{-1} R^k e_\ell = 0$$

and thus Ψ is not positive definite. □

The above lemma allows us to write down the best linear unbiased estimator.

Definition 4.8 (Best linear unbiased estimator)**.** If Ψ is well defined and positive definite we define the *best linear unbiased estimator* (BLUE)

$$\widehat{\mu}_\alpha^{\mathrm{B}} := \sum_{k \in U_S} m_k \alpha^T \Psi^{-1} P^k (C^k)^{-1} \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right).$$

This is a linear unbiased estimator with coefficients

$$\beta^k := m_k (C^k)^{-1} R^k \Psi^{-1} \alpha. \tag{4.15}$$

◇

The BLUE is the linear unbiased estimator for $\alpha^T \mu$ with the smallest variance. This justifies its name.

Theorem 4.9 (Existence and uniqueness of BLUE)**.** Let the covariance matrices C^k be positive definite for all $k \in U_S$ and assume that we evaluate every model at least once $U_Z = \{1, \dots, L\}$. Then the estimator $\widehat{\mu}_\alpha^{\mathrm{B}}$ is a linear unbiased estimator for $\alpha^T \mu$ with variance

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \alpha^T \Psi^{-1} \alpha. \tag{4.16}$$

Every linear unbiased estimator $\widehat{\mu}_\alpha$ that uses the same samples as $\widehat{\mu}_\alpha^{\mathrm{B}}$ satisfies

$$\mathbb{V}[\widehat{\mu}_\alpha] \geq \mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big]$$

and equality holds only if $\widehat{\mu}_\alpha = \widehat{\mu}_\alpha^{\mathrm{B}}$.

Proof. We verify that the coefficients in (4.15) are the unique best choice. First, we write down the convex minimization problem for minimizing the variance such that the bias constraint is satisfied

$$\begin{aligned} \min_{\beta} \qquad & J(\beta) := \sum_{k \in U_S} \frac{(\beta^k)^T C^k \beta^k}{m_k} \\ \text{such that} \qquad & \alpha = \sum_{k \in U_S} P^k \beta^k. \end{aligned} \tag{4.17}$$

We evaluate every model $U_Z = \{1, \dots, L\}$ and thus there exists a linear unbiased estimator which is a feasible point of (4.17). The KKT conditions are thus necessary and sufficient for a minimizer. The KKT conditions with the Lagrange-multiplier $\lambda \in \mathbb{R}^L$ read

$$\frac{1}{m_k} C^k \beta^k + R^k \lambda = 0 \qquad \text{for all } k \in U_S, \tag{4.18}$$

$$\sum_{k \in U_S} P^k \beta^k = \alpha. \tag{4.19}$$

We solve for β^k in (4.18) since C^k is positive and insert the result into (4.19) to obtain an expression for λ

$$\alpha = -\sum_{k \in U_S} m_k P^k (C^k)^{-1} R^k \lambda = -\Psi \lambda.$$

Since Ψ is invertible we solve for λ and insert the result in (4.18)

$$\frac{1}{m_k} C^k \beta^k - R^k \Psi^{-1} \alpha = 0 \qquad \text{for all } k \in U_S.$$

Again we solve for β^k and obtain the coefficients (4.15) which are the minimizer of (4.17). Since these coefficients satisfy the bias constraint (4.19) the BLUE is unbiased $\mathbb{E}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \alpha$. We insert the coefficients into the variance expression in (4.13)

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \sum_{k \in U_S} \frac{(\beta^k)^T C^k \beta^k}{m_k} = \sum_{k \in U_S} m_k \alpha^T \Psi^{-1} P^k (C^k)^{-1} R^k \Psi^{-1} \alpha = \alpha^T \Psi^{-1} \Psi \Psi^{-1} \alpha,$$

which is the variance (4.16). Now let $\widehat{\mu}_\alpha$ be another linear unbiased estimator. Then to minimize the variance we are forced to choose the coefficients to be uniform for every sample in each model group. The resulting estimator then corresponds to a feasible point in (4.17) where $\widehat{\mu}_\alpha^{\mathrm{B}}$ is the unique minimizer. Thus $\mathbb{V}[\widehat{\mu}_\alpha] > \mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big]$ if $\widehat{\mu}_\alpha \neq \widehat{\mu}_\alpha^{\mathrm{B}}$. □

Remark 4.10 (Existence and uniqueness of BLUE for $U_Z \subsetneq \{1, \dots, L\}$)**.** The condition $U_Z = \{1, \dots, L\}$ is not necessary to derive a BLUE. As long as there exists a linear unbiased estimator $U_\alpha \subseteq U_Z$ we are able to construct the BLUE by removing the unused models $\{1, \dots, L\} \setminus U_Z$ and renumbering the remaining QoIs such that $Z_1, \dots, Z_{L'}$ with $U_Z = \{1, \dots, L'\}$. Then we apply Theorem 4.9 to this smaller set where every model is used. ◇

Linear models. In the previous paragraph we derived the BLUE from a constructive perspective. Fix the bias α, ensure that a linear unbiased estimator exists and select the coefficients β such that the variance is minimized. This follows the control variate approach in Section 3.3. We now derive this estimator from the perspective of linear models and linear regression. The key observation is to reformulate the estimation and to interpret the random variable Z_L as noisy observation of the unknown mean μ_L in the sense of a linear model

$$Z_L = \mu_L + (Z_L - \mu_L) = \mu_L + \eta_L.$$

Here η_L is a mean zero noise with variance σ_L^2. We extend this basic idea for all QoIs $Z_1, \dots, Z_L$, for each model group and for multiple i.i.d. samples of every model group. Recall that we want to estimate the vector $\mu \in \mathbb{R}^L$ or some linear combination $\alpha^T \mu$ of it.

Definition 4.11 (Linear model)**.** For $k \in \{1, \dots, K\}$ we abbreviate

$$\begin{aligned} Z_{S^k} &:= (Z_\ell)_{\ell \in S^k}, \\ \eta_{S^k} &:= Z_{S^k} - R^k \mu, \end{aligned}$$

and define the linear model

$$Z_{S^k} = R^k \mu + \eta_{S^k}.$$

We collect m_k samples of Z_{S^k} and define Y^k, B^k and η^k such that

$$Y^k := \begin{pmatrix} Z_{S^k}^1 \\ \vdots \\ Z_{S^k}^{m_k} \end{pmatrix} = \begin{pmatrix} R^k \\ \vdots \\ R^k \end{pmatrix} \mu + \begin{pmatrix} \eta_{S^k}^1 \\ \vdots \\ \eta_{S^k}^{m_k} \end{pmatrix} = B^k \mu + \eta^k.$$

We again collect these vectors in a *linear model* such that with Y, B and η it holds

$$Y := \begin{pmatrix} Y^1 \\ \vdots \\ Y^K \end{pmatrix} = \begin{pmatrix} B^1 \\ \vdots \\ B^K \end{pmatrix} \mu + \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^K \end{pmatrix} = B\mu + \eta. \tag{4.20}$$

◇

In the context of linear models Y is called the vector of observations, B the design matrix, μ the parameter that we want to estimate and η the noise vector. We give a brief example with three models to explicitly express such a linear model.

Example 4.12 (Linear model)**.** Let $L := 3$ and the $K = 7$ model groups be

$$\begin{aligned} &S^1 := \{1\}, && S^2 := \{2\}, && S^3 := \{3\}, && S^4 := \{1, 2\}, \\ &S^5 := \{1, 3\}, && S^6 := \{2, 3\}, && S^7 := \{1, 2, 3\}. \end{aligned}$$

We define the number of samples as $m_1 := m_7 := 1, m_6 := 2$ and $m_k := 0$ for all other model groups. Ignoring the number of samples for a moment, the linear models are

$$\begin{aligned} Z_1 &= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + (Z_1 - \mu_1) && = R^1 \mu + \eta_{\{1\}}, \\ \begin{pmatrix} Z_2 \\ Z_3 \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \left(\begin{pmatrix} Z_2 \\ Z_3 \end{pmatrix} - \begin{pmatrix} \mu_2 \\ \mu_3 \end{pmatrix} \right) && = R^6 \mu + \eta_{\{2,3\}}, \\ \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \left(\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} \right) && = R^7 \mu + \eta_{\{1,2,3\}}. \end{aligned}$$

We now collect these in a block vector to account for two samples of S^6. We arrive at

$$\left(\begin{array}{c} Z_1^{1,1} \\ \hline Z_2^{1,6} \\ Z_3^{1,6} \\ \hline Z_2^{2,6} \\ Z_3^{2,6} \\ \hline Z_1^{1,7} \\ Z_2^{1,7} \\ Z_3^{1,7} \end{array}\right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \left(\begin{array}{c} Z_1^{1,1} - \mu_1 \\ \hline Z_2^{1,6} - \mu_2 \\ Z_3^{1,6} - \mu_3 \\ \hline Z_2^{2,6} - \mu_2 \\ Z_3^{2,6} - \mu_3 \\ \hline Z_1^{1,7} - \mu_1 \\ Z_2^{1,7} - \mu_2 \\ Z_3^{1,7} - \mu_3 \end{array}\right) = \begin{pmatrix} B^1 \\ B^6 \\ B^7 \end{pmatrix} \mu + \begin{pmatrix} \eta^1 \\ \eta^6 \\ \eta^7 \end{pmatrix} = B\mu + \eta.$$

◇

Linear models are well studied in mathematics and we derive some properties of this specific instance (4.20). The design matrix $B \in \mathbb{R}^{N \times L}$ is typically a tall and skinny matrix, since we often have more samples N than levels L in the following sense

$$N := \sum_{k=1}^{K} m_k |S^k| \gg L.$$

Similarly, realizations of the observations and noise $Y(\omega), \eta(\omega) \in \mathbb{R}^N$ are vectors of size N. The covariance matrix of the noise η describes how much the observations deviate from the mean μ and the correlation between different observations. The noise covariance matrix of a model group is equal to the model group covariance matrix

$$\mathbb{C}\mathrm{ov}[\eta_{S^k}, \eta_{S^k}] = \mathbb{C}\mathrm{ov}[Z_{S^k}, Z_{S^k}] = P^k \mathbb{C}\mathrm{ov}\big[Z_{\{1,\dots,L\}}, Z_{\{1,\dots,L\}}\big] R^k = P^k C R^k = C^k.$$

For the special case of i.i.d. samples across model groups and samples in the same model group the covariance matrices of the combined noise vectors η^k and η are block diagonal.

Lemma 4.13 (Mean and covariance of η)**.** The random variables η_{S^k}, η^k and η have mean zero. For all $k \in \{1, \dots, K\}$ with i.i.d. samples $Z_{S^k}^1, \dots, Z_{S^k}^{m_k}$ the covariance matrix of the noise η^k is block diagonal

$$\mathbb{C}\mathrm{ov}\big[\eta^k, \eta^k\big] = \mathrm{diag}\big((C^k)_{i=1}^{m_k}\big). \tag{4.21}$$

For i.i.d. samples also across model groups $Z_{S^k}^{1,k}, \dots, Z_{S^k}^{m_k,k}$ for all $k \in \{1, \dots, K\}$ the covariance matrix of the noise η is also block diagonal

$$\mathbb{C}\mathrm{ov}[\eta, \eta] = \mathrm{diag}\big((\mathbb{C}\mathrm{ov}\big[\eta^k, \eta^k\big])_{k=1}^K\big) = \mathrm{diag}\big(((C^k)_{i=1}^{m_k})_{k=1}^K\big). \tag{4.22}$$

Proof. The proof that the noises η_{S^k}, η^k and η have mean zero follows directly from Definition 4.11. We now show (4.21). We denote with $[i]$ the entry of η^k that corresponds to the i-th sample $\eta^{[i],k} := Z^i_{S^k} - \mu_{S^k}$. For $i, j \in \{1, \dots, m_k\}$ the independence across the sample index then shows the block diagonal form

$$\mathbb{C}\mathrm{ov}\big[\eta^{[i],k}, \eta^{[j],k}\big] = \mathbb{C}\mathrm{ov}\big[Z^i_{S^k}, Z^j_{S^k}\big] = \begin{cases} C^k, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Similarly, for (4.22) we denote with $[i,k]$ the entry of η that correspond to the i-th sample for the k-th model group $\eta^{[i,k]} := Z^{i,k}_{S^k} - \mu_{S^k}$. For $k, \ell \in \{1, \dots, K\}$ and $i \in \{1, \dots, m_k\}$, $j \in \{1, \dots, m_\ell\}$ the independence across the sample index and model group shows the block diagonal form

$$\mathbb{C}\mathrm{ov}\big[\eta^{[i,k]}, \eta^{[j,\ell]}\big] = \mathbb{C}\mathrm{ov}\Big[Z^{i,k}_{S^k}, Z^{j,\ell}_{S^\ell}\Big] = \begin{cases} C^k \in \mathbb{R}^{|S^k| \times |S^k|}, & \text{if } i = j \text{ and } k = \ell, \\ 0 \in \mathbb{R}^{|S^k| \times |S^\ell|}, & \text{otherwise.} \end{cases}$$ □

We are now in a position to derive the BLUE from the perspective of a generalized linear least squares problem. First, recall the linear model in (4.20)

$$Y = B\mu + \eta.$$

We model η as mean zero Gaussian with covariance $\mathbb{C}\mathrm{ov}[\eta, \eta]$ and thus

$$Y \sim N(B\mu, \mathbb{C}\mathrm{ov}[\eta, \eta]).$$

This distribution is determined solely by the parameter μ, since we assume that C and thus $\mathbb{C}\mathrm{ov}[\eta, \eta]$ is known. The goal is now to compute the the *maximum likelihood estimator* for μ, which is the point where the probability density function is maximized. This is equivalent to minimizing the negative log likelihood

$$\begin{aligned} \widehat{\mu}^{\mathrm{B}} &:= \operatorname*{argmax}_{\mu \in \mathbb{R}^L} \quad c \exp\left(-\frac{1}{2} (B\mu - Y)^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} (B\mu - Y) \right) \\ &= \operatorname*{argmin}_{\mu \in \mathbb{R}^L} \quad \frac{1}{2} (B\mu - Y)^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} (B\mu - Y). \end{aligned} \tag{4.23}$$

The estimator $\widehat{\mu}^{\mathrm{B}}$ in (4.23) is the solution of a least squares problem where the norm is determined by the symmetric, positive definite and block diagonal matrix $\mathbb{C}\mathrm{ov}[\eta, \eta]$. We set the derivative of the objective function to zero such that the minimizer $\widehat{\mu}^{\mathrm{B}}$ satisfies the normal equations

$$B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} B \widehat{\mu}^{\mathrm{B}} = B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} Y. \tag{4.24}$$

Solving the normal equations means inverting the matrix on the left–hand side of (4.24) to arrive at

$$\widehat{\mu}^{\mathrm{B}} = (B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} B)^{-1} B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} Y. \tag{4.25}$$

We now verify that this estimator is the BLUE for estimating $\widehat{\mu}$. We start by simplifying the expressions and show that the system matrix in (4.24) is equal to Ψ.

Lemma 4.14 (Least squares solution)**.** Let C^k be positive definite for all $k \in U_S$ and assume that we evaluate every model $U_Z = \{1, \dots, L\}$. Then the estimator $\widehat{\mu}^{\mathrm{B}}$ in (4.25) is well defined and

$$\Psi = B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} B = \sum_{k \in U_S} m_k P^k (C^k)^{-1} R^k, \tag{4.26}$$

$$\widehat{\mu}^{\mathrm{B}} = \Psi^{-1} \sum_{k \in U_S} m_k P^k (C^k)^{-1} \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right). \tag{4.27}$$

Proof. We use the block diagonal structure of the covariance matrix of η in (4.22) and the invertibility of C^k to conclude that $\mathbb{C}\mathrm{ov}[\eta, \eta]$ is invertible. We further write down the inverse and recall the expression for B and Y

$$\begin{aligned} \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} &= \mathrm{diag}\big((([C^k]^{-1})_{i=1}^{m_k})_{k \in U_S} \big), \\ B &= ((R^k)_{i=1}^{m_k})_{k \in U_S}, \\ Y &= ((Z_{S^k}^{i,k})_{i=1}^{m_k})_{k \in U_S}. \end{aligned}$$

A calculation now shows the expression for the matrix Ψ in (4.26)

$$B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} B = B^T (([C^k]^{-1} R^k)_{i=1}^{m_k})_{k \in U_S} = \sum_{k \in U_S} \sum_{i=1}^{m_k} P^k (C^k)^{-1} R^k = \Psi.$$

The right–hand side of the normal equations (4.24) is

$$B^T \mathbb{C}\mathrm{ov}[\eta, \eta]^{-1} Y = B^T (([C^k]^{-1} Z_{S^k}^{i,k})_{i=1}^{m_k})_{k \in U_S} = \sum_{k \in U_S} \sum_{i=1}^{m_k} P^k (C^k)^{-1} Z_{S^k}^{i,k}.$$

We combine this result with (4.25), (4.26) and the invertibility of Ψ in the proof of Theorem 4.9 to conclude (4.27). □

The previous lemma allows us to show that the estimator $\widehat{\mu}^{\mathrm{B}}$ is actually the BLUE for estimating the mean μ.

Theorem 4.15 (Gauss–Markov–Aitken [114, Theorem 4.4])**.** Let C^k be positive definite for all $k \in U_S$ and assume that we evaluate every model $U_Z = \{1, \dots, L\}$. Then the estimator $\widehat{\mu}^{\mathrm{B}}$ is a linear unbiased estimator of μ with covariance

$$\mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}}, \widehat{\mu}^{\mathrm{B}}\big] = \Psi^{-1}. \tag{4.28}$$

Every linear unbiased estimator $\widehat{\mu}$ that uses the same samples satisfies

$$\mathbb{C}\mathrm{ov}[\widehat{\mu}, \widehat{\mu}] \geq \mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}}, \widehat{\mu}^{\mathrm{B}}\big] \tag{4.29}$$

and equality holds only if $\widehat{\mu} = \widehat{\mu}^{\mathrm{B}}$.

Proof. The proof is a translation of [114, Section 4.2] or [76, Appendix A] for our setting. We first show the unbiasedness of the BLUE

$$\mathbb{E}\big[\widehat{\mu}^{\mathrm{B}}\big] = \Psi^{-1} \sum_{k \in U_S} m_k P^k (C^k)^{-1} \left(\frac{1}{m_k} \sum_{i=1}^{m_k} \mathbb{E}\Big[Z_{S^k}^{i,k} \Big] \right) = \Psi^{-1} \sum_{k \in U_S} m_k P^k (C^k)^{-1} R^k \mu = \mu.$$

We use (4.27) to express the covariance matrix of the BLUE as

$$\mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\widehat{\mu}^{\mathrm{B}}\big] = \Psi^{-1}\mathbb{C}\mathrm{ov}\left[\sum_{k\in U_S} P^k(C^k)^{-1}\sum_{i=1}^{m_k} Z^{i,k}_{S^k}, \sum_{k\in U_S} P^k(C^k)^{-1}\sum_{i=1}^{m_k} Z^{i,k}_{S^k}\right]\Psi^{-1}.$$

The independence across the sample index i and model groups k now shows (4.28)

$$\begin{aligned}\mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\widehat{\mu}^{\mathrm{B}}\big] &= \Psi^{-1}\sum_{k\in U_S}\sum_{i=1}^{m_k} P^k(C^k)^{-1}\mathbb{C}\mathrm{ov}\Big[Z^{i,k}_{S^k}, Z^{i,k}_{S^k}\Big](C^k)^{-1}R^k\Psi^{-1}\\ &= \Psi^{-1}\sum_{k\in U_S} m_k P^k(C^k)^{-1}C^k(C^k)^{-1}R^k\Psi^{-1}\\ &= \Psi^{-1}\Psi\Psi^{-1}.\end{aligned}$$

Now let $\widehat{\mu}$ be another linear estimator for μ that uses exactly the same samples

$$\widehat{\mu} = \widehat{\mu}^{\mathrm{B}} + \beta^T Y,$$

where $\beta \in \mathbb{R}^{N\times L}$ models the respective difference for the linear combination compared to the BLUE. We require that $\widehat{\mu}$ is unbiased for all possible mean vectors $\mu \in \mathbb{R}^L$

$$\mu = \mathbb{E}[\widehat{\mu}] = \mathbb{E}\big[\widehat{\mu}^{\mathrm{B}} + \beta^T Y\big] = \mu + \beta^T B\mu + \beta^T\mathbb{E}[\eta] = \mu + \beta^T B\mu,$$

from which we conclude $\beta^T B = 0$. The covariance matrix of $\widehat{\mu}$ is

$$\mathbb{C}\mathrm{ov}[\widehat{\mu},\widehat{\mu}] = \mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\widehat{\mu}^{\mathrm{B}}\big] + 2\mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\beta^T Y\big] + \beta^T\mathbb{C}\mathrm{ov}[Y,Y]\beta.$$

We use $Y = B\mu + \eta$ and $\beta^T B = 0$ to conclude that the second term is zero

$$\begin{aligned}\mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\beta^T Y\big] &= \mathbb{C}\mathrm{ov}\big[\Psi^{-1}B^T\mathbb{C}\mathrm{ov}[\eta,\eta]^{-1}Y,\beta^T Y\big]\\ &= \mathbb{C}\mathrm{ov}\big[\Psi^{-1}B^T\mathbb{C}\mathrm{ov}[\eta,\eta]^{-1}\eta,\beta^T\eta\big]\\ &= \Psi^{-1}B^T\mathbb{C}\mathrm{ov}[\eta,\eta]^{-1}\mathbb{C}\mathrm{ov}[\eta,\eta]\beta\\ &= \Psi^{-1}B^T\beta\\ &= 0.\end{aligned}$$

We conclude (4.28) since the matrix $\mathbb{C}\mathrm{ov}[Y,Y] = \mathbb{C}\mathrm{ov}[\eta,\eta]$ is positive definite and thus for a positive semi–definite matrix A

$$\mathbb{C}\mathrm{ov}[\widehat{\mu},\widehat{\mu}] = \mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\widehat{\mu}^{\mathrm{B}}\big] + \beta^T\mathbb{C}\mathrm{ov}[\eta,\eta]\beta = \mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}},\widehat{\mu}^{\mathrm{B}}\big] + A.$$

We now verify that $A = 0$ implies $\beta = 0$ and therefore $\widehat{\mu} = \widehat{\mu}^{\mathrm{B}}$ is the unique BLUE. The covariance matrix of the noise is block diagonal, hence the matrix A satisfies

$$A = \beta^T\mathbb{C}\mathrm{ov}[\eta,\eta]\beta = \beta^T((C^k\beta^{i,k})_{i=1}^{m_k})_{k\in U_S} = \sum_{k\in U_S}\sum_{i=1}^{m_k}(\beta^{i,k})^T C^k\beta^{i,k},$$

where $\beta^{i,k}\in\mathbb{R}^{|S^k|\times L}$ are the respective entries of β. Notice that the matrix $(\beta^{i,k})^T C^k\beta^{i,k} \in \mathbb{R}^{L\times L}$ is positive semi–definite and thus $A = 0$ implies that for all $v\in\mathbb{R}^L$

$$v^T(\beta^{i,k})^T C^k\beta^{i,k}v = 0.$$

We use that C^k is positive definite to conclude

$$\beta^{i,k}v = 0.$$

Since this result holds for all vectors $v\in\mathbb{R}^L$ we conclude $\beta^{i,k} = 0$ and thus $\beta = 0$. □

We remark that the BLUE for the specific linear combination $\alpha^T\mu$ is given by

$$\widehat{\mu}_\alpha^{\mathrm{B}} = \alpha^T \widehat{\mu}^{\mathrm{B}}$$

and thus we immediately obtain an expression for its variance

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \mathbb{C}\mathrm{ov}\big[\widehat{\mu}_\alpha^{\mathrm{B}}, \widehat{\mu}_\alpha^{\mathrm{B}}\big] = \alpha^T \mathbb{C}\mathrm{ov}\big[\widehat{\mu}^{\mathrm{B}}, \widehat{\mu}^{\mathrm{B}}\big]\alpha = \alpha^T \Psi^{-1}\alpha.$$

We furthermore obtain the coefficients

$$\beta^k = m_k (C^k)^{-1} R^k \Psi^{-1} \alpha.$$

This coincides with the BLUE in Definition 4.8 and Theorem 4.9 is now a straightforward consequence of Theorem 4.15.

4.2 Lower variance bound

We derive the lower variance bound for the BLUE and thus for all linear unbiased estimators. We assume that we evaluate the model group S^1 which contains the high fidelity model once and all other model groups with models in $I^\infty \subseteq \{1, \dots, L-1\}$ an arbitrary amount of time. The fixed models $I^F := S^1 \setminus I^\infty$ are then evaluated only once and the models that reduce the variance are $I := I^\infty \cap S^1$. We assume that $U_Z = \{1, \dots, L\} = S^1 \cup I^\infty$, otherwise some models are not used and can w.l.o.g. be removed. This setting is a slight generalization of the control variate setting where $S^1 = \{1, \dots, L\}$, $I^\infty = I = \{1, \dots, L-1\}$ and $I^F = \{L\}$. We verify that the BLUE has the same lower variance bound as the CV or ACV estimators. Let us give a brief example for the aforementioned sets.

Example 4.16 (Lower variance bound sets)**.** Let $S^1 := \{1, 3, 4\}$ and assume that we have an arbitrary amount of model evaluations in $I^\infty := \{2, 3, 5, 6\}$. We want to compute the lower variance bound for the BLUE with bias $\alpha := (1, 1, 1, 1, 0, 0)^T$. We then have

$$\begin{aligned} U_Z &= S^1 \cup I^\infty = \{1, \dots, 6\}, \\ I^F &= \{1, 4\}, \\ I &= \{3\}. \end{aligned}$$

◇

We will later show that the coefficients of models in I^F are fixed and only samples in I can be used to reduce the variance. The lower variance bound for the BLUE is the solution of a minimization problem.

Lemma 4.17 (Lower variance bound optimization)**.** The lower variance bound of the BLUE is

$$\mathbb{V}^{\min}[\widehat{\mu}_\alpha^{\mathrm{B}}] = \min_\beta \quad \beta^T C_{I,I}\beta - 2\beta^T C_{I,I^F}\alpha_{I^F} + \alpha_{I^F}^T C_{I^F,I^F}\alpha_{I^F}. \tag{4.30}$$

Proof. "≥": We write down the variance of the BLUE in terms of a minimization problem

$$\begin{aligned} \mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \min_\beta \quad & (\beta^1)^T C^1 \beta^1 + \sum_{k \in U_S \setminus \{1\}} \frac{(\beta^k)^T C^k \beta^k}{m_k} \\ \text{such that} \quad & \alpha = \sum_{k \in U_S} P^k \beta^k. \end{aligned} \tag{4.31}$$

The entries of β^1 corresponding to I^F are fixed, since these models are evaluated only once and only in S^1

$$\alpha_{I^F} = \sum_{k \in U_S} \beta^k_{[I^F]} = \beta^1_{[I^F]}.$$

We drop the remaining bias constraints and the variance terms not belonging to S^1. This yields a lower bound for the variance

$$\mathbb{V}\big[\widehat{\mu}^{\mathrm{B}}_\alpha\big] \geq \min_{\beta^1_{[I]}} \quad \begin{pmatrix} \beta^1_{[I]} \\ \beta^1_{[I^F]} \end{pmatrix}^T \begin{pmatrix} C_{I,I} & C_{I,I^F} \\ C_{I^F,I} & C_{I^F,I^F} \end{pmatrix} \begin{pmatrix} \beta^1_{[I]} \\ \beta^1_{[I^F]} \end{pmatrix}.$$

We rename $\beta^1_{[I^F]} = \alpha_{I^F}$ and $\beta^1_{[I]} = -\beta$ to obtain exactly the lower bound in (4.30).
"$\leq$": Let β^* be a minimizer of (4.30) and $\beta^1_{[I^F]} = \alpha_{I^F}$. We insert these suboptimal coefficients in (4.31) to obtain the upper bound

$$\mathbb{V}\big[\widehat{\mu}^{\mathrm{B}}_\alpha\big] \leq \begin{pmatrix} \beta^* \\ \alpha_{I^F} \end{pmatrix}^T \begin{pmatrix} C_{I,I} & C_{I,I^F} \\ C_{I^F,I} & C_{I^F,I^F} \end{pmatrix} \begin{pmatrix} \beta^* \\ \alpha_{I^F} \end{pmatrix} + \min_{\beta^k} \sum_{k \in U_S \setminus \{1\}} \frac{(\beta^k)^T C^k \beta^k}{m_k},$$

where we still have to satisfy the bias constraint with the other β^k for $k \neq 1$. This constraint can always be achieved by samples of the model group $S^{k^\infty} := I^{k^\infty}$

$$\alpha = \begin{pmatrix} \alpha_{I^{k^\infty}} \\ \alpha_{I^F} \end{pmatrix} = P^1 \begin{pmatrix} \beta^* \\ \alpha_{I^F} \end{pmatrix} + P^{k^\infty} \beta^{k^\infty} + \sum_{k \in U_S \setminus \{1, k^\infty\}} P^k \beta^k$$

if we set $\beta^k := 0$ for $k \in U_S \setminus \{1, k^\infty\}$, $\beta^{k^\infty}_{[I]} := \alpha_I - \beta^*$ and $\beta^{k^\infty}_{[I^{k^\infty} \setminus I]} := \alpha_{I^{k^\infty} \setminus I}$. The variance is thus upper bounded

$$\mathbb{V}\big[\widehat{\mu}^{\mathrm{B}}_\alpha\big] \leq \begin{pmatrix} \beta^* \\ \alpha_{I^F} \end{pmatrix}^T \begin{pmatrix} C_{I,I} & C_{I,I^F} \\ C_{I^F,I} & C_{I^F,I^F} \end{pmatrix} \begin{pmatrix} \beta^* \\ \alpha_{I^F} \end{pmatrix} + \frac{(\beta^{k^\infty})^T C^{k^\infty} \beta^{k^\infty}}{m_{k^\infty}}.$$

No coefficient here depends on the number of samples m_{k^∞}, hence the limit procedure for $m_{k^\infty} \to +\infty$ and afterwards using the definition of β^* shows the result. □

The lower variance bound is a Schur complement and for special sets and biases equal to the bound for the control variate setting.

Corollary 4.18 (Lower variance bound for the BLUE)**.** For positive definite $C_{I,I}$ the lower variance bound of the BLUE is a Schur complement

$$\mathbb{V}^{\min}[\widehat{\mu}^{\mathrm{B}}_\alpha] = \alpha^T_{I^F} \big(C_{I^F,I^F} - C_{I^F,I} C^{-1}_{I,I} C_{I,I^F}\big) \alpha_{I^F}.$$

For $\alpha := e_L$, $I := I^\infty := \{1, \dots, L-1\}$ and $S^1 := \{1, \dots, L\}$ this bound is equal to the lower variance bound for the control variates

$$\mathbb{V}^{\min}[\widehat{\mu}^{\mathrm{B}}_L] = \mathbb{V}^{\min}[\widehat{\mu}^{\mathrm{CV}}_L] = \mathbb{V}^{\min}_I = C \setminus C_{I,I}.$$

Proof. The minimizer in (4.30) satisfies

$$C_{I,I} \beta = C_{I,I^F} \alpha_{I^F},$$

where β can be computed since $C_{I,I}$ is invertible. The proof of the remainder of this corollary is then a straightforward computation. □

We summarize the result of this section in the next theorem, which states a tight lower bound for the variance of any linear unbiased estimator with a certain sample allocation.

Theorem 4.19 (Lower variance bound)**.** Assume that $C_{I,I}$ is positive definite. Let $\widehat{\mu}_L$ be a linear unbiased estimator that uses a single sample of $Z^1_{S^1}$ and further samples of Z^i_ℓ with $\ell \in I^\infty$ which are uncorrelated to Z_{S^1}. Then its variance is lower bounded

$$\mathbb{V}[\widehat{\mu}_L] \geq \mathbb{V}_I^{\min}. \tag{4.32}$$

A linear unbiased estimator $\widehat{\mu}$ for μ with the same samples as $\widehat{\mu}_L$ satisfies

$$\mathbb{C}\mathrm{ov}[\widehat{\mu}, \widehat{\mu}]_{I^F,I^F} \geq C_{I^F,I^F} - C_{I^F,I} C_{I,I}^{-1} C_{I,I^F}. \tag{4.33}$$

Both inequalities are tight.

Proof. We prove both inequalities (4.32) and (4.33) by replacing $\widehat{\mu}_L$ with the BLUE that uses the same sample allocation. By definition of the BLUE this never increases the variance

$$\mathbb{V}[\widehat{\mu}_L] \geq \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big].$$

We now use $\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] \geq \mathbb{V}_I^{\min}$ from Corollary 4.18 to show (4.32). The tightness of this bound follows if $\widehat{\mu}_L$ is the BLUE with a suitable sample allocation. We obtain (4.33) and its tightness by observing that for all $\alpha \in \mathbb{R}^L$ with $\alpha = (0, \alpha_{I^F})^T$

$$\begin{aligned}\alpha_{I^F}^T \mathbb{C}\mathrm{ov}[\widehat{\mu}, \widehat{\mu}]_{I^F,I^F} \alpha_{I^F} &= \alpha^T \mathbb{C}\mathrm{ov}[\widehat{\mu}, \widehat{\mu}]\alpha = \mathbb{V}\big[\alpha^T \widehat{\mu}\big] \geq \mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] \\ &\geq \alpha_{I^F}^T \big(C_{I^F,I^F} - C_{I^F,I} C_{I,I}^{-1} C_{I,I^F}\big)\alpha_{I^F}.\end{aligned}$$

□

Up until this point we made the choice to evaluate the sample group S^1. Importantly, all models that are evaluated only once are in this model group. We now explain this choice with an example.

Example 4.20 (Sample allocation)**.** Let us assume that the models in I^∞ and the models in S^1 are uncorrelated, that is

$$\mathbb{C}\mathrm{ov}[Z_\ell, Z_j] = 0, \quad \text{for all } \ell \in I^\infty, j \in S^1.$$

Then clearly, no variance reduction can be achieved and we are allowed to ignore the bias constraints associated with these models. Thus w.l.o.g. $I^\infty = \emptyset$. We further assume that we have two models and $\alpha_1 = \alpha_2 = 1$. Then there exists a unique linear unbiased estimator which is equal to the BLUE with variance

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \mathbb{V}[\alpha_1 Z_1 + \alpha_2 Z_2].$$

Notice that if we restrict the number of evaluations of Z_1 and Z_2 we are free to choose whether we use dependent samples and a single evaluation of $S^1 := \{1, 2\}$ or independent samples with a single evaluation of both $S^2 := \{1\}$ and $S^3 := \{2\}$. The variance of the two different BLUEs is then

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \begin{cases} \sigma_1^2 + \sigma_2^2, & \text{for } m_1 = 1, m_2 = 0, m_3 = 0, \\ \sigma_1^2 + \sigma_2^2 + 2C_{1,2}, & \text{for } m_1 = 0, m_2 = 1, m_3 = 1. \end{cases}$$

Clearly, depending on the sign of $C_{1,2}$ either of these estimators may achieve a smaller variance than the other. ◇

We conclude that we have two different BLUEs that achieve a different variance. The smallest variance is achieved if we also optimize over the possible sample allocations. We deal with this problem in Chapter 5.

4.3 Linear subspace formulation

Lower variance bound. Up until this point we formulated the BLUE as a specific linear combination

$$\widehat{\mu}_L^{\mathrm{B}} = \sum_{k \in U_S} (\beta^k)^T \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right)$$

and we chose the coefficients to minimize the variance such that a bias constraint is satisfied. This approach focuses on the coefficients β and not the linear subspace spanned by the linear combination of the MC estimators. In this rather short section, we interpret the BLUE with the help of the best approximation in a certain Hilbert space similarly to [60], where this was done with control variates. It is well known that least square solutions can often be interpreted as best approximations, or equivalently, that computing a best approximation requires the solution of a least squares problem. We define the subspace of L^2 with mean zero random variables

$$H := \{Z - \mathbb{E}[Z] \,|\, Z \in L^2\}$$

with the inner product

$$(Z, Y) := \mathbb{C}\mathrm{ov}[Z, Y].$$

This space is a well–defined Hilbert space since the scalar product is positive definite on H due to the restriction to mean zero random variables. Formally, for $Z \in L^2$ with

$$(Z, Z) = \mathbb{C}\mathrm{ov}[Z, Z] = \mathbb{V}[Z] = 0$$

we conclude that $Z = \mathbb{E}[Z]$ is almost surely constant. The assumption $Z \in H$ then shows $Z = \mathbb{E}[Z] = 0$ almost surely.
For an index set $I \subseteq \{1, \dots, L-1\}$ we define the subspace of control variates

$$V_I := \operatorname{span}\big((Z_\ell - \mu_\ell)_{\ell \in I}\big) \subseteq H.$$

The lower variance bound in Lemma 4.17 is thus a minimization problem over this space

$$\begin{aligned}
\mathbb{V}_I^{\min} &= \min_\beta \ \beta^T C_{I,I} \beta - 2\beta^T C_{I,L} + \sigma_L^2 \\
&= \min_\beta \begin{pmatrix} -\beta \\ 1 \end{pmatrix}^T \begin{pmatrix} C_{I,I} & C_{I,L} \\ C_{L,I} & \sigma_L^2 \end{pmatrix} \begin{pmatrix} -\beta \\ 1 \end{pmatrix} \\
&= \min_\beta \mathbb{V}\left[Z_L - \sum_{\ell \in I} \beta_{[\ell]} Z_\ell \right] \\
&= \min_{v \in V_I} \|Z_L - v\|_H^2 \\
&= \operatorname{dist}(Z_L, V_I).
\end{aligned}$$

The minimizer β^* now corresponds to the element $v^* \in V_I$ which is the best approximation of $Z_L - \mu_L \in H$. The residual is orthogonal to all elements in V_I

$$Z_L - \mu_L - v^* \perp V_I,$$

in other words the residual is uncorrelated to every element in the space V_I. Therefore, the best approximation v^* is the orthogonal projection of $Z_L - \mu_L$ onto the space V_I. We remark that the CV estimator assumes that $\mu_1, \dots, \mu_{L-1}$ is known and thus the space V_I is known. The best approximation is the linear combination of the control variates that minimize the variances. This perspective is often helpful as we show in the next example.

Example 4.21 (Linear combination)**.** Let Z_L be a linear combination of $Z_1, \dots, Z_{L-1}$ and let ξ denote a random variable which is uncorrelated to $Z_1, \dots, Z_{L-1}$

$$Z_L := \sum_{\ell=1}^{L-1} \beta_\ell Z_\ell + \xi, \tag{4.34}$$

then ξ is the residual and its variance is the lower variance bound

$$\mathbb{V}^{\min}_{\{1,\dots,L-1\}} = \mathbb{V}[\xi].$$

Conversely, the expansion (4.34) for Z_L always exists and if $C_{I,I}$ with $I := \{1, \dots, L-1\}$ is positive definite, then the coefficients β_ℓ are unique. ◇

Best approximation for multiple samples. Let us now write down the BLUE for multiple model groups. We restrict ourselves to the estimator with $S^L := \{1, \dots, L\}$ and non–zero m_L, which allows us to easily express the results. We first eliminate the bias constraint by fixing β^L as

$$\beta^L = \sum_{k \in U_S \setminus \{L\}} P^k \beta^k - e_L.$$

We insert this expression into the variance (4.13) of linear unbiased estimators and with $I := \{1, \dots, L-1\}$ we obtain

$$\begin{aligned}
\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] =& \frac{(\beta^L)^T C \beta^L}{m_L} + \sum_{k \in U_S \setminus \{L\}} \frac{(\beta^k)^T C^k \beta^k}{m_k} \\
=& \left(\sum_{k \in U_S \setminus \{L\}} (\beta^k)^T R^k \right) \frac{C}{m_L} \left(\sum_{k \in U_S \setminus \{L\}} P^k \beta^k \right) - 2 \left(\sum_{k \in U_S \setminus \{L\}} (\beta^k)^T R^k \right) \frac{C_{I,L}}{m_L} + \frac{\sigma_L^2}{m_L} \\
&+ \sum_{k \in U_S \setminus \{L\}} \frac{(\beta^k)^T C^k \beta^k}{m_k}.
\end{aligned}$$

A computation shows that the above expression is actually equal to

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] = \mathbb{V}\left[\frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^{i,L} - \sum_{k \in U_S \setminus \{L\}} (\beta^k)^T \left(\frac{1}{m_L} \sum_{i=1}^{m_L} Z_{S^k}^{i,L} - \frac{1}{m_k} \sum_{i=1}^{m_k} Z_{S^k}^{i,k} \right) \right].$$

We define the control variate space, that now depends on the number of samples m

$$V_I^m := \operatorname{span}\left(\left(\left(\frac{1}{m_L} \sum_{i=1}^{m_L} Z_\ell^{i,L} - \frac{1}{m_k} \sum_{i=1}^{m_k} Z_\ell^{i,k} \right)_{\ell \in S^k} \right)_{k \in U_S \setminus \{L\}} \right) \subseteq H.$$

The basis functions are differences of uncorrelated MC estimators which ensures that the estimator remains unbiased. Then the best approximation minimizes the distance to the MC estimator of Z_L in the following sense

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] = \min_{v \in V_I^m} \left\| \frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^{i,L} - v \right\|_H^2 = \operatorname{dist}\left(\frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^{i,L}, V_I^m \right).$$

Example 4.22 (Linear combination)**.** Let the MC estimator of Z_L be a linear combination of the functions that define V_I^m plus some random variable ξ with $\xi \perp V_I^m$

$$\begin{aligned}\frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^{i,L} &:= \sum_{k\in U_S}\sum_{\ell\in S^k} \beta_{[\ell]}^k \left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} - \frac{1}{m_k}\sum_{i=1}^{m_k} Z_\ell^{i,k}\right) + \xi \\ &= \sum_{k\in U_S}\sum_{\ell\in S^k} \beta_{[\ell]}^k \left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L}\right) - \sum_{k\in U_S}\sum_{\ell\in S^k} \beta_{[\ell]}^k \left(\frac{1}{m_k}\sum_{i=1}^{m_k} Z_\ell^{i,k}\right) + \xi.\end{aligned}$$

It is always possible to find such a decomposition. We see that we now have two parts of the estimator that depend on the coefficients β, where the first part is correlated with the MC estimator of Z_L and the second part is not. The variance of the BLUE is

$$\begin{aligned}\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] &= \mathbb{V}\left[\frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^{i,L} - \sum_{k\in U_S}\sum_{\ell\in S^k} \beta_{[\ell]}^k \frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L}\right] + \mathbb{V}\left[\sum_{k\in U_S}\sum_{\ell\in S^k} \beta_{[\ell]}^k \frac{1}{m_k}\sum_{i=1}^{m_k} Z_\ell^{i,k}\right] \\ &= \mathbb{V}[\xi].\end{aligned}$$

◇

4.4 Comparison of linear unbiased estimators

In this section we derive the circumstances under which the linear unbiased estimators of Chapter 3, that is the MC, CV, ACV, MFMC and MLMC estimators, are the BLUE or give examples when they are not. These estimators are all of the form

$$\widehat{\mu}_L = \sum_{i=1}^{N} \beta_i Z_{\ell_i}^{k_i} \tag{4.35}$$

for suitable coefficients β_i. We further write down the model groups to denote the correlation structure and thus the underlying linear regression problem. In this section, we assume that the covariance matrix C is positive definite to ensure that the coefficients β for e.g. the BLUE or the CV estimator are well defined and unique.

Monte Carlo. We first look at the MC estimator

$$\widehat{\mu}_L^{\mathrm{MC}} := \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^i$$

which uses a single model group

$$S^L = \{L\}.$$

We use the convention that the model Z_L is in the samples group S^L and we keep this consistent for this section.
The MC estimator is the BLUE. A formal verification uses Lemma 4.6 to show that the coefficients β_i in (4.35) for the BLUE are all equal and thus $\beta_i = 1/m_L$, which shows that the BLUE is the MC estimator. If $\sigma_L^2 = 0$ and thus C is not positive definite, then the BLUE is not unique and any combination of coefficients with

$$\sum_{i=1}^{N} \beta_i = 1$$

delivers a BLUE with zero variance.

Control Variates. We deal with the CV estimator in a special way, since these estimators assume that $\mu_1, \dots, \mu_{L-1}$ is known

$$\widehat{\mu}_L^{\mathrm{CV}} := \frac{1}{m_L} \sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=1}^{L-1} \beta_\ell \left(\frac{1}{m_L} \sum_{i=1}^{m_L} Z_\ell^i - \mu_\ell \right).$$

The only model group that is used consists of all models

$$S^L = \{1, \dots, L\}.$$

The knowledge of $\mu_1, \dots, \mu_{L-1}$ is in stark contrast to the other estimators and to the BLUE discussed in this chapter. Therefore, we compare it to a hypothetical linear unbiased estimator

$$\widehat{\mu}_L := \sum_{i=1}^{m_L} (\beta^i)^T Z_{S^L}^i$$

where we only keep the bias constraint for the fine model Z_L

$$1 = \sum_{i=1}^{m_L} \beta_L^i.$$

This can be achieved since we assume that $\mu_1, \dots, \mu_{L-1}$ are known. The variance of this estimator is

$$\mathbb{V}[\widehat{\mu}_L] = \sum_{i=1}^{m_L} (\beta^i)^T C \beta^i.$$

The KKT conditions for optimizing the coefficients β^i and adhering to the bias constraint require us to find a Lagrange multiplier $\lambda \in \mathbb{R}$ such that

$$\begin{aligned} C\beta^i + \lambda e_L &= 0 \qquad \text{for all } i \in \{1, \dots, m_L\}, \\ \sum_{i=1}^{m_L} e_L^T \beta^i &= 1. \end{aligned}$$

A straightforward calculation shows that there exists a unique solution such that all coefficients are equal w.r.t.the sample index i

$$\beta^i = \frac{1}{m_L} C^{-1} e_L (e_L^T C^{-1} e_L)^{-1} = \frac{1}{m_L} \begin{pmatrix} -\beta \\ 1 \end{pmatrix}$$

for a suitably chosen β that minimizes the variance. Hence the BLUE is the CV estimator

$$\widehat{\mu}_L^{\mathrm{B}} = \sum_{i=1}^{m_L} (\beta^i)^T Z_{S^L}^i = \frac{1}{m_L} \sum_{i=1}^{m_L} \left(Z_L^i - \sum_{\ell=1}^{L-1} \beta_\ell Z_\ell^i \right) = \widehat{\mu}_L^{\mathrm{CV}}.$$

For positive definite covariance matrix C the choice of β^i is unique. This shows existence and uniqueness of the BLUE and the optimality of the CV estimator for the special case of knowing $\mu_1, \dots, \mu_{L-1}$ with $S^L = \{1, \dots, L\}$.

Multifidelity Monte Carlo. The MFMC estimator assumes that the expectations $\mu_1, \dots, \mu_{L-1}$ are unknown and have to be estimated. This estimator is

$$\widehat{\mu}_L^{\mathrm{MFMC}} := \frac{1}{n_L}\sum_{i=1}^{n_L} Z_L^i - \sum_{\ell=1}^{L-1}\beta_\ell\left(\frac{1}{n_{\ell+1}}\sum_{i=1}^{n_{\ell+1}} Z_\ell^i - \frac{1}{n_\ell}\sum_{i=1}^{n_\ell} Z_\ell^i\right).$$

The model groups for this approach are

$$\begin{aligned} S^1 &= \{1\}, & m_1 &= n_1 - n_2, \\ &\vdots & &\vdots \\ S^{L-1} &= \{1, \dots, L-1\}, & m_{L-1} &= n_{L-1} - n_L, \\ S^L &= \{1, \dots, L\}, & m_L &= n_L. \end{aligned}$$

We verify that this estimator is the BLUE for some interesting cases.

Theorem 4.23 (Cases for which MFMC is BLUE)**.** Let C be positive definite. Assume that at least one of the following is true

(i) $L \leq 2$,

(ii) Z_{L-1} and Z_L are uncorrelated to $Z_1, \dots, Z_{L-2}$, that is $\rho_{L,\ell} = \rho_{L-1,\ell} = 0$ for $\ell \in \{1, \dots, L-2\}$.

Then the MFMC estimator is the BLUE $\widehat{\mu}_L^{\mathrm{MFMC}} = \widehat{\mu}_L^{\mathrm{B}}$.

Proof. "(i)": For $L = 1$ the MFMC estimator is the MC estimator and thus the BLUE. For $L = 2$ the bias constraint of the BLUE ensures that $\beta_2^2 = 1$ and thus

$$\widehat{\mu}_L^{\mathrm{B}} = \frac{1}{m_2}\sum_{i=1}^{m_2}(\beta^2)^T\begin{pmatrix} Z_1^{i,2} \\ Z_2^{i,2}\end{pmatrix} + \frac{1}{m_1}\sum_{i=1}^{m_1}\beta_1^1 Z_1^{i,1} = \frac{1}{m_2}\sum_{i=1}^{m_2} Z_2^{i,2} + \beta_1^2\frac{1}{m_2}\sum_{i=1}^{m_2} Z_1^{i,2} + \beta_1^1\frac{1}{m_1}\sum_{i=1}^{m_1} Z_1^{i,1}.$$

A computation redefining the coefficients β_1^2 and β_1^1 in a suitable manner shows

$$\beta_1^2\frac{1}{m_2}\sum_{i=1}^{m_2} Z_1^{i,2} + \beta_1^1\frac{1}{m_1}\sum_{i=1}^{m_1} Z_1^{i,1} = \widetilde{\beta}_1^2\frac{1}{m_2}\sum_{i=1}^{m_2} Z_1^{i,2} + \widetilde{\beta}_1^1\frac{1}{m_1+m_2}\left(\sum_{i=1}^{m_1} Z_1^{i,1} + \sum_{i=1}^{m_2} Z_1^{i,2}\right).$$

The bias constraint now requires that $\widetilde{\beta}_1^2 = -\widetilde{\beta}_1^1$ and thus the BLUE is equal to the MFMC estimator

$$\widehat{\mu}_L^{\mathrm{B}} = \frac{1}{m_2}\sum_{i=1}^{m_2} Z_2^{i,2} - \widetilde{\beta}_1^1\left(\frac{1}{m_2}\sum_{i=1}^{m_2} Z_1^{i,2} - \frac{1}{m_1+m_2}\left(\sum_{i=1}^{m_1} Z_1^{i,1} + \sum_{i=1}^{m_2} Z_1^{i,2}\right)\right) = \widehat{\mu}_L^{\mathrm{MFMC}}.$$

This result follows, since both estimators optimize over the coefficient $\widetilde{\beta}_1^1$.
"(ii)": For $L > 2$ with $\rho_{L,\ell} = 0$ for $\ell \in \{1, \dots, L-2\}$ the coefficients of the MFMC estimator from (3.22) are equal to zero

$$\beta_\ell = \frac{C_{L\ell}}{\sigma_\ell^2} = \rho_{L\ell}\frac{(\sigma_L^2\sigma_\ell^2)^{1/2}}{\sigma_\ell^2} = 0.$$

Therefore, this estimator is equal to the MFMC estimator with $L = 2$ after renaming Z_{L-1}, Z_L to Z_1, Z_2. We prove that the BLUE and the MFMC estimator are equal if we

verify that the BLUE also only uses Z_{L-1} and Z_L. First, define $I := \{L-1, L\}$ and with $I^c = \{1, \dots, L-2\}$ we use $C_{I,I^c} = 0$ to conclude that the covariance matrix is block diagonal. A computation then also shows $\Psi_{I,I^c} = 0$ and thus Ψ is block diagonal. We combine this with the formula for the coefficients (4.15) of the BLUE

$$\begin{aligned}
\beta^k &= m_k (C^k)^{-1} R^k \Psi^{-1} e_L \\
&= m_k \begin{pmatrix} C^{-1}_{I^c \cap S^k, I^c \cap S^k} & 0 \\ 0 & C^{-1}_{I \cap S^k, I \cap S^k} \end{pmatrix} R^k \begin{pmatrix} \Psi^{-1}_{I^c, I^c} & 0 \\ 0 & \Psi^{-1}_{I,I} \end{pmatrix} \begin{pmatrix} 0_{|I^c|} \\ 0 \\ 1 \end{pmatrix} \\
&= m_k \begin{pmatrix} C^{-1}_{I^c \cap S^k, I^c \cap S^k} & 0 \\ 0 & C^{-1}_{I \cap S^k, I \cap S^k} \end{pmatrix} R^k \begin{pmatrix} 0_{|I^c|} \\ \Psi^{-1}_{I,I} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \\
&= m_k R^k \begin{pmatrix} C^{-1}_{I^c \cap S^k, I^c \cap S^k} & 0 & 0 \\ 0 & C^{-1}_{I \cap S^k, I \cap S^k} & 0 \\ 0 & 0 & 0_{|(S^k)^c|} \end{pmatrix} \begin{pmatrix} 0_{|I^c|} \\ \Psi^{-1}_{I,I} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix}.
\end{aligned}$$

A careful inspection using $S^k = \{1, \dots, k\}$ and $I^c = \{1, \dots, L-2\}$ now shows that $\beta^k_{[I^c]} = 0$ and thus the BLUE does not use $Z_1, \dots, Z_{L-2}$. □

We comment on the different cases. Case (i) says that the MFMC estimator is optimal if we have two models and case (ii) says that this is still the case if we add models that are uncorrelated to Z_{L-1} and Z_L. For $L > 2$ the MFMC estimator is typically not the BLUE and we provide and example for $L = 3$. We exploit the property that the MFMC estimator does not depend on the correlations $\rho_{\ell j}$ for models $j, \ell \in \{1, \dots, L-1\}$.

Example 4.24 (Noisy observations)**.** We define the models such that

$$\begin{aligned}
Z_1 &:= Y + 10\xi_1 + \xi_2, \\
Z_2 &:= 10Y + 10\xi_1, \\
Z_3 &:= 10Y,
\end{aligned}$$

where $Y, \xi_1, \xi_2 \sim N(0,1)$ are independent random variables. Notice that the coarsest model Z_1 is almost independent of Z_3 but can be used to remove the noise ξ_1 from Z_2. Therefore the BLUE leverages the linear combination $Z_3 \approx Z_2 - Z_1$ which yields a small variance. The MFMC estimator is not able to do this. We write down the covariance and correlation matrices

$$C = \begin{pmatrix} 102 & 110 & 10 \\ 110 & 200 & 100 \\ 10 & 100 & 100 \end{pmatrix}, \qquad \mathrm{P} \approx \begin{pmatrix} 1 & 0.77 & 0.10 \\ 0.77 & 1 & 0.71 \\ 0.10 & 0.71 & 1 \end{pmatrix}.$$

We conclude that the ordering of the models satisfies $0 < \rho_{1,3} < \rho_{2,3} < \rho_{3,3} = 1$ which is required by the MFMC Theorem 3.26. The variance of the MFMC estimator is

$$\begin{aligned}
\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] &= \sigma_L^2 \sum_{\ell=1}^{L} \frac{\rho^2_{L,\ell} - \rho^2_{L,\ell-1}}{n_\ell} \geq \frac{100}{n_3}(1 - 1/2) + \frac{100}{n_2}(1/2 - 0.01) + \frac{100}{n_1} 0.01 \\
&= \frac{50}{n_3} + \frac{49}{n_2} + \frac{1}{n_1}.
\end{aligned}$$

Contrast this with the following MLMC estimator that uses fewer samples

$$\widehat{\mu}_L^{\mathrm{MLMC}} = \frac{1}{n_3}\sum_{i=1}^{n_3}(Z_3^{i,3} - (Z_2^{i,3} - Z_1^{i,3})) + \frac{1}{n_2 - n_3}\sum_{i=1}^{n_2-n_3}(Z_2^{i,2} - Z_1^{i,2}).$$

This is a linear unbiased estimator and thus the variance of the BLUE satisfies

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] \leq \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MLMC}}\big] = \frac{2}{n_3} + \frac{82}{n_2 - n_3}.$$

A suitable choice of n_2 and n_3 now shows that the MFMC estimator is not the BLUE

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] \leq \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MLMC}}\big] < \mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big]. \qquad \diamond$$

Let us continue with another interesting example. The MFMC estimator only uses the correlations $\rho_{\ell,L}$ for $\ell \in \{1,\dots,L\}$ and thus it is tempting to think that this estimator is the BLUE if $\rho_{\ell,j} = 0$ for $\ell, j \in \{1,\dots,L-1\}$ with $\ell \neq j$, that is, no correlation between the models Z_ℓ and Z_j can be used. This is however, not the case. We derive this from the lower variance bound.

Example 4.25 (MFMC is not BLUE)**.** We define the QoI as follows

$$\begin{aligned} Z_1 &:= \xi_1. \\ Z_2 &:= \xi_2, \\ Z_3 &:= Y + \xi_1 + \xi_2, \end{aligned}$$

where again $Y, \xi_1, \xi_2 \in N(0,1)$ are independent random variables. Clearly Z_1 and Z_2 are uncorrelated. However, the lower variance bound for the BLUE is tight and thus there exists n_1, n_2, n_3 and $\varepsilon > 0$ such that

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{B}}\big] \leq (1+\varepsilon)\mathbb{V}^{\min}[\widehat{\mu}_L^{\mathrm{B}}] = (1+\varepsilon)\mathbb{V}^{\min}_{\{1,2\}} = (1+\varepsilon)\mathbb{V}[Y] = 1+\varepsilon.$$

On the other hand, we have shown that the MFMC estimator only reaches the lower variance bound in Theorem 3.34

$$\mathbb{V}\big[\widehat{\mu}_L^{\mathrm{MFMC}}\big] \geq \mathbb{V}^{\min}[\widehat{\mu}_L^{\mathrm{MFMC}}] = \mathbb{V}^{\min}_{\{2\}} = \mathbb{V}[Y + \xi_1] = 2.$$

Therefore, the MFMC estimator is not the BLUE. $\diamond$

Approximate Control Variates. We start with the ACV–IS estimator which is defined as follows

$$\widehat{\mu}_L^{\text{ACV–IS}} := \frac{1}{n_L}\sum_{i=1}^{n_L} Z_L^{i,L} - \sum_{\ell=1}^{L-1}\beta_\ell\left(\frac{1}{n_L}\sum_{i=1}^{n_L} Z_\ell^{i,L} - \frac{1}{n_\ell}\left(\sum_{i=1}^{n_L} Z_\ell^{i,L} + \sum_{i=n_L+1}^{n_\ell} Z_\ell^{i,\ell}\right)\right).$$

The sample groups and their respective number of samples are

$$\begin{aligned} S^1 &= \{1\}, & m_1 &= n_1 - n_L, \\ &\vdots & &\vdots \\ S^{L-2} &= \{L-2\}, & m_{L-2} &= n_{L-2} - n_L, \\ S^{L-1} &= \{L-1\}, & m_{L-1} &= n_{L-1} - n_L, \\ S^L &= \{1,\dots,L\}, & m_L &= n_L. \end{aligned}$$

It turns out that the ACV–IS estimator is the BLUE. The intuitive reasoning is that the BLUE with the same sample allocation requires the coefficients to satisfy $\beta_{[\ell]}^{\ell} = -\beta_{\ell}^{L}$ from the bias constraint. Then the BLUE is essentially a control variate approach where the unknown $\mu_1, \dots, \mu_{L-1}$ are estimated from independent MC estimators from samples of the model groups $S^1, \dots, S^{L-1}$.

Theorem 4.26 (ACV–IS is BLUE)**.** Let C be positive definite. Then the ACV–IS estimator is the BLUE, $\widehat{\mu}_L^{\text{ACV–IS}} = \widehat{\mu}_L^{\text{B}}$.

Proof. First, we rewrite the ACV–IS estimator to group the respective model groups

$$\begin{aligned}\widehat{\mu}_L^{\text{ACV–IS}} &= \frac{1}{n_L}\sum_{i=1}^{n_L} Z_L^{i,L} - \sum_{\ell=1}^{L-1}\beta_\ell\left(\left(\frac{1}{n_L}-\frac{1}{n_\ell}\right)\sum_{i=1}^{n_L} Z_\ell^{i,L} - \frac{1}{n_\ell}\sum_{i=n_L+1}^{n_\ell} Z_\ell^{i,\ell}\right)\\ &= \frac{1}{n_L}\sum_{i=1}^{n_L} Z_L^{i,L} - \sum_{\ell=1}^{L-1}\beta_\ell\frac{m_\ell}{n_\ell}\left(\frac{n_\ell}{m_\ell}\left(\frac{1}{n_L}-\frac{1}{n_\ell}\right)\sum_{i=1}^{n_L} Z_\ell^{i,L} - \frac{1}{m_\ell}\sum_{i=n_L+1}^{n_\ell} Z_\ell^{i,\ell}\right)\end{aligned}$$

We use $m_\ell = n_\ell - n_L$ for $\ell \in \{1, \dots, L-1\}$ and $n_L = m_L$ to conclude that

$$\frac{n_\ell}{m_\ell}\left(\frac{1}{n_L}-\frac{1}{n_\ell}\right) = \frac{n_\ell}{m_\ell}\frac{n_\ell - n_L}{n_L n_\ell} = \frac{1}{n_L} = \frac{1}{m_L}.$$

We change the superscript of the samples $Z_\ell^{i,\ell}$ and the definition of β_ℓ to arrive at

$$\widehat{\mu}_L^{\text{ACV–IS}} = \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^{i,L} - \sum_{\ell=1}^{L-1}\widetilde{\beta}_\ell\left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^{i,\ell}\right). \tag{4.36}$$

The BLUE with the same sample allocation is

$$\widehat{\mu}_L^{\text{B}} = \frac{1}{m_L}\sum_{i=1}^{m_L}(\beta^L)^T Z_{S^L}^{i,L} + \sum_{\ell=1}^{L-1}\beta_1^\ell\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^{i,\ell}.$$

Here the bias constraint requires the coefficients to satisfy

$$\begin{aligned}&\beta_\ell^L = -\beta_{[\ell]}^\ell = -\beta_1^\ell \qquad \text{for all } \ell \in \{1, \dots, L-1\},\\ &\beta_L^L = 1.\end{aligned}$$

We use this result and group the models Z_ℓ together

$$\widehat{\mu}_L^{\text{B}} = \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^{i,L} - \sum_{\ell=1}^{L-1}\beta_1^\ell\left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^{i,\ell}\right).$$

This however, is exactly (4.36) for $\widetilde{\beta}_\ell = \beta_1^\ell$. The ACV–IS estimator and the BLUE are identical $\widehat{\mu}_L^{\text{ACV–IS}} = \widehat{\mu}_L^{\text{B}}$, since both select the coefficients to minimize the variance. □

We continue with the ACV–MF estimator

$$\widehat{\mu}_L^{\text{ACV–MF}} := \frac{1}{n_L}\sum_{i=1}^{n_L} Z_L^{i} - \sum_{\ell=1}^{L-1}\beta_\ell\left(\frac{1}{n_L}\sum_{i=1}^{n_L} Z_\ell^{i} - \frac{1}{n_\ell}\sum_{i=1}^{n_\ell} Z_\ell^{i}\right), \tag{4.37}$$

where we w.l.o.g. assume that $n_{L-1} \leq n_{L-2} \leq \cdots \leq n_1$, otherwise we rename $Z_1, \ldots, Z_{L-1}$ accordingly. The model groups and number of samples are the same as for the MFMC estimator

$$\begin{aligned} S^1 &= \{1\}, & m_1 &= n_1 - n_2, \\ &\vdots & &\vdots \\ S^{L-1} &= \{1, \ldots, L-1\}, & m_{L-1} &= n_{L-1} - n_L, \\ S^L &= \{1, \ldots, L\}, & m_L &= n_L. \end{aligned}$$

This estimator is the MFMC estimator if $L \leq 2$ or if $Z_1, \ldots, Z_{L-2}$ are uncorrelated to Z_{L-1}, Z_L and thus this estimator is the BLUE in this case.

Theorem 4.27 (Cases for which ACV–MF is BLUE)**.** Let C be positive definite. Assume that at least one of the following is true

(i) $L \leq 2$,

(ii) Z_{L-1} and Z_L are uncorrelated to $Z_1, \ldots, Z_{L-2}$, that is $\rho_{L,\ell} = \rho_{L-1,\ell} = 0$ for $\ell \in \{1, \ldots, L-2\}$,

(iii) $Z_1, \ldots, Z_{L-1}$ are pairwise uncorrelated.

Then the ACV–MF estimator is the BLUE, $\widehat{\mu}_L^{\text{ACV–MF}} = \widehat{\mu}_L^{\text{B}}$.

Proof. "(i)": The estimators are equal $\widehat{\mu}_L^{\text{MFMC}} = \widehat{\mu}_L^{\text{ACV–MF}}$ and Theorem 4.23 now shows the result.
"(ii)": We reduce this case to (i) by showing that only the two fine grid models are used. With $I := \{1, \ldots, L-1\}$ and $J := \{1, \ldots, L-2\}$ from Lemma 3.39 we use the block diagonal structure of the covariance matrix such that for all $\ell \in \{1, \ldots, L-2\}$

$$\begin{aligned} \beta_\ell &= e_\ell^T (F^{MF} \circ C_{I,I})^{-1} (\operatorname{diag}(F^{MF}) \circ C_{I,L}) \\ &= e_\ell^T \begin{pmatrix} (F^{MF} \circ C_{J,J})^{-1} & 0 \\ 0 & (F^{MF}_{L-1,L-1} C_{L-1,L-1})^{-1} \end{pmatrix} \begin{pmatrix} 0 \\ F^{MF}_{L-1,L-1} C_{L-1,L} \end{pmatrix} \\ &= e_\ell^T \begin{pmatrix} 0 \\ \frac{1}{C_{L-1,L-1}} C_{L-1,L} \end{pmatrix} \\ &= 0. \end{aligned}$$

Therefore the ACV–MF estimator only uses the models Z_{L-1} and Z_L an we may w.l.o.g. assume that $L = 2$. Now (i) shows the result.
"(iii)": We show that the BLUE is equal to the ACV–MF estimator. The coefficients of the BLUE divided by the number of samples are constant for $k \in \{1, \ldots, L-1\}$ and $\ell \in \{1, \ldots, k\}$. This follows from the diagonal structure of $C^k = \operatorname{diag}((\sigma_\ell^2)_{\ell=1}^k)$

$$\frac{\beta_\ell^k}{m_k} = e_\ell^T (C^k)^{-1} R^k \Psi^{-1} e_L = \frac{1}{\sigma_\ell^2} e_\ell^T R^k \Psi^{-1} e_L = \frac{1}{\sigma_\ell^2} e_\ell^T \Psi^{-1} e_L =: \widetilde{\beta}_\ell.$$

This coefficient is independent of k and only depends on ℓ. Therefore, the BLUE is

$$\begin{aligned}\widehat{\mu}_L^{\mathrm{B}} &= \sum_{k=1}^{L}\sum_{\ell=1}^{k}\frac{\beta_\ell^k}{m_k}\sum_{i=1}^{m_\ell} Z_\ell^{i,k} \\ &= (\beta^L)^T\frac{1}{m_L}\sum_{i=1}^{m_L} Z_{S^L}^{i,L} + \sum_{k=1}^{L-1}\sum_{\ell=1}^{L-1}\widetilde{\beta}_\ell 1_{S^k}(\ell)\sum_{i=1}^{m_k} Z_\ell^{i,k} \\ &= (\beta^L)^T\frac{1}{m_L}\sum_{i=1}^{m_L} Z_{S^L}^{i,L} + \sum_{\ell=1}^{L-1}\widetilde{\beta}_\ell\sum_{k=1}^{L-1}\sum_{i=1}^{m_k} 1_{S^k}(\ell) Z_\ell^{i,k}.\end{aligned}$$

Let us now collect all samples that evaluate the ℓ-th model for $\ell \in \{1,\dots,L-1\}$ and introduce the coefficients $\check{\beta}_\ell, \overline{\beta}_\ell$

$$\begin{aligned}&\beta_\ell^L\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} + \widetilde{\beta}_\ell\sum_{k=1}^{L-1}\sum_{i=1}^{m_k} 1_{S^k}(\ell) Z_\ell^{i,k} \\ &= \left(\frac{\beta_\ell^L}{m_L} - \widetilde{\beta}_\ell\right)\sum_{i=1}^{m_L} Z_\ell^{i,L} + \widetilde{\beta}_\ell\sum_{k=1}^{L}\sum_{i=1}^{m_k} 1_{S^k}(\ell) Z_\ell^{i,k} \\ &= \check{\beta}_\ell\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} + \overline{\beta}_\ell\frac{1}{\sum_{j=\ell}^{L} m_j}\sum_{k=1}^{L}\sum_{i=1}^{m_k} 1_{S^k}(\ell) Z_\ell^{i,k}.\end{aligned}$$

Notice that the right expression is a MC estimator with $\sum_{j=\ell}^{L} m_j = n_\ell$ independent samples. We rewrite this and use the bias constraint $\check{\beta}_\ell = -\overline{\beta}_\ell$ to obtain

$$\check{\beta}_\ell\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^{i,L} + \overline{\beta}_\ell\frac{1}{\sum_{j=\ell}^{L} m_j}\sum_{k=1}^{L}\sum_{i=1}^{m_k} 1_{S^k}(\ell) Z_\ell^{i,k} = -\overline{\beta}_\ell\left(\frac{1}{n_L}\sum_{i=1}^{n_L} Z_\ell^i - \frac{1}{n_\ell}\sum_{i=1}^{n_\ell} Z_\ell^i\right).$$

The bias constraint is satisfied only if $\beta_L^L = 1$ and the combination of these results shows

$$\widehat{\mu}_L^{\mathrm{B}} = \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=1}^{L-1}\overline{\beta}_\ell\left(\frac{1}{n_L}\sum_{i=1}^{n_L} Z_\ell^i - \frac{1}{n_\ell}\sum_{i=1}^{n_\ell} Z_\ell^i\right).$$

With $m_L = n_L$ this is exactly the ACV–MF estimator in (4.37). The statement now follows since both estimators choose $\overline{\beta}_\ell$ such that the variance is minimized. □

We give an example that the ACV–MF estimator is not the BLUE for $L \geq 3$ in the numerical experiments in Section 4.5. Let us continue with the ACV–KL estimator

$$\begin{aligned}\widehat{\mu}_L^{\text{ACV–KL}} :=& \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^i - \sum_{\ell=K}^{L-1}\beta_\ell\left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_\ell^i - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^i\right) \\ &- \sum_{\ell=1}^{K-1}\beta_\ell\left(\frac{1}{m_N}\sum_{i=1}^{m_N} Z_\ell^i - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_\ell^i\right).\end{aligned}$$

The model groups and number of samples are the same as for the MFMC and ACV–MF estimator

$$\begin{aligned} S^1 &= \{1\}, & m_1 &= n_1 - n_2, \\ &\vdots & &\vdots \\ S^{L-1} &= \{1,\dots,L-1\}, & m_{L-1} &= n_{L-1} - n_L, \\ S^L &= \{1,\dots,L\}, & m_L &= n_L. \end{aligned}$$

Since the ACV–KL estimator is equal to the ACV–MF estimator for $N = L$ or $K = 1$ and these parameters are chosen to minimize the variance, we use Theorem 4.27 to obtain the following result.

Theorem 4.28 (Cases for which ACV–KL is BLUE)**.** Let C be positive definite. Assume that at least one of the following is true

(i) $L \leq 2$,

(ii) Z_{L-1} and Z_L are uncorrelated to $Z_1, \dots, Z_{L-2}$, that is $\rho_{L,\ell} = \rho_{L-1,\ell} = 0$ for $\ell \in \{1, \dots, L-2\}$,

(iii) $Z_1, \dots, Z_{L-1}$ are uncorrelated.

Then the ACV–KL estimator is the BLUE $\widehat{\mu}_L^{\text{ACV–KL}} = \widehat{\mu}_L^{\text{B}}$ and $N = L$ or $K = 1$. ◇

We postpone the verification that the ACV–KL estimator is in general not the BLUE for $L \geq 3$ to the numerical experiments in Section 4.5. We only verify this for some fixed values of N or K by using the lower variance bound.

Example 4.29 (ACV–KL is not BLUE)**.** The lower variance bound of the BLUE and of the ACV–KL estimator in Theorem 3.43 with $N < L$ satisfies

$$\mathbb{V}^{\min}[\widehat{\mu}^{\text{ACV–KL}}] = \mathbb{V}^{\min}_{\{K,\dots,L-1\}} \geq \mathbb{V}^{\min}_{\{1,\dots,L-1\}} = \mathbb{V}^{\min}[\widehat{\mu}_L^{\text{B}}].$$

The above inequality is strict for suitably defined $Z_1, \dots, Z_L$ and K. As an example, assume that $Z_1, \dots, Z_{L-1}$ are pairwise independent and have unit variance, that is $\mathbb{V}[Z_\ell] := 1$ for all $\ell \in \{1, \dots, L-1\}$. Now define the high fidelity model as follows

$$Z_L = \sum_{\ell=1}^{L-1} Z_\ell + \xi,$$

where $\xi \sim N(0,1)$ is independent of $Z_1, \dots, Z_{L-1}$. Then the inequality is strict if $K < L$

$$L - K + 1 = \mathbb{V}^{\min}_{\{K,\dots,L-1\}} > \mathbb{V}^{\min}_{\{1,\dots,L-1\}} = 1.$$ ◇

Multilevel Monte Carlo. The MLMC estimator is defined as

$$\widehat{\mu}_L^{\text{MLMC}} := \sum_{\ell=1}^{L} \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} (Z_\ell^{i,\ell} - Z_{\ell-1}^{i,\ell}).$$

Here we defined $Z_0 := 0$ and thus the sample groups are

$$\begin{aligned} S^1 &= \{1\}, \\ S^2 &= \{1, 2\}, \\ &\vdots \\ S^{L-1} &= \{L-2, L-1\}, \\ S^L &= \{L-1, L\}. \end{aligned} \tag{4.38}$$

It is straightforward to verify that this estimator is not the BLUE. The reason is that we do not optimize the coefficients in front of the samples. In particular, scaling $Z_1, \dots, Z_{L-1}$ by a non–zero constant does not change the variance of the BLUE but changes the variance of the MLMC estimator. Furthermore, the MLMC estimator is not able to ignore models that are independent of Z_L by setting the coefficients in front of them to zero.

Example 4.30 (MLMC is not BLUE)**.** We choose $Z_1, Z_2 \sim N(0,1)$ to be independent random variables. Then

$$\mathbb{V}\big[\widehat{\mu}_2^{\mathrm{MLMC}}\big] = \frac{2}{m_2} + \frac{1}{m_1}.$$

Contrast this with the BLUE that only uses the last model

$$\mathbb{V}\big[\widehat{\mu}_2^{\mathrm{B}}\big] = \frac{1}{m_2}.$$

We conclude that the MLMC estimator is not the BLUE

$$\mathbb{V}\big[\widehat{\mu}_2^{\mathrm{MLMC}}\big] > \mathbb{V}\big[\widehat{\mu}_2^{\mathrm{B}}\big]. \qquad \diamond$$

Full coupling estimator. For completeness we define the *full coupling* (FC) estimator which we use in the numerical experiments Section 4.5. This estimator uses the identical sample allocation as the MFMC, ACV–MF and ACV–KL estimators

$$\begin{aligned} S^1 &:= \{1\}, \\ &\vdots \\ S^{L-1} &:= \{1, \dots, L-1\}, \\ S^L &:= \{1, \dots, L\} \end{aligned}$$

and is defined as the associated BLUE

$$\widehat{\mu}_L^{\mathrm{FC}} := \widehat{\mu}_L^{\mathrm{B}}.$$

We further the define FC κ estimator that couples only the last κ models

$$\begin{aligned} S^{L-\kappa+1} &:= \{L-\kappa+1\}, \\ &\vdots \\ S^{L-1} &:= \{L-\kappa+1, \dots, L-1\}, \\ S^L &:= \{L-\kappa+1, \dots, L\} \end{aligned}$$

and is again the BLUE with this sample allocation

$$\widehat{\mu}_L^{\mathrm{FC}\,\kappa} := \widehat{\mu}_L^{\mathrm{B}}. \tag{4.39}$$

We clearly have $\widehat{\mu}_L^{\mathrm{FC}\,L} = \widehat{\mu}_L^{\mathrm{FC}}$ and that $\widehat{\mu}_L^{\mathrm{FC}\,1}$ is the MC estimator. A calculation similar to Section 4.3 to eliminate the bias constraint shows that this estimator is very similar to a CV estimator

$$\widehat{\mu}_L^{\mathrm{FC}\,\kappa} := \frac{1}{m_L}\sum_{i=1}^{m_L} Z_L^{i,L} - \sum_{\ell=L-\kappa+1}^{L-1} (\beta^\ell)^T \left(\frac{1}{m_L}\sum_{i=1}^{m_L} Z_{S^\ell}^{i,L} - \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_{S^\ell}^{i,\ell} \right).$$

The coefficients $\beta^\ell \in \mathbb{R}^{|S^\ell|}$ are once again chosen such that the variance is minimized.

Summary. Let us summarize the used model groups for every estimator of this section in Figure 4.1. We further summarize under which circumstances the estimators are the BLUE in Table 4.1. We define the conditions

$$L \leq 2 \quad \text{or} \quad \rho_{L,\ell} = \rho_{L-1,\ell} = 0 \text{ for all } \ell \in \{1, \dots, L-2\}, \tag{4.40}$$

$$Z_1, \dots, Z_{L-1} \text{ are pairwise uncorrelated.} \tag{4.41}$$

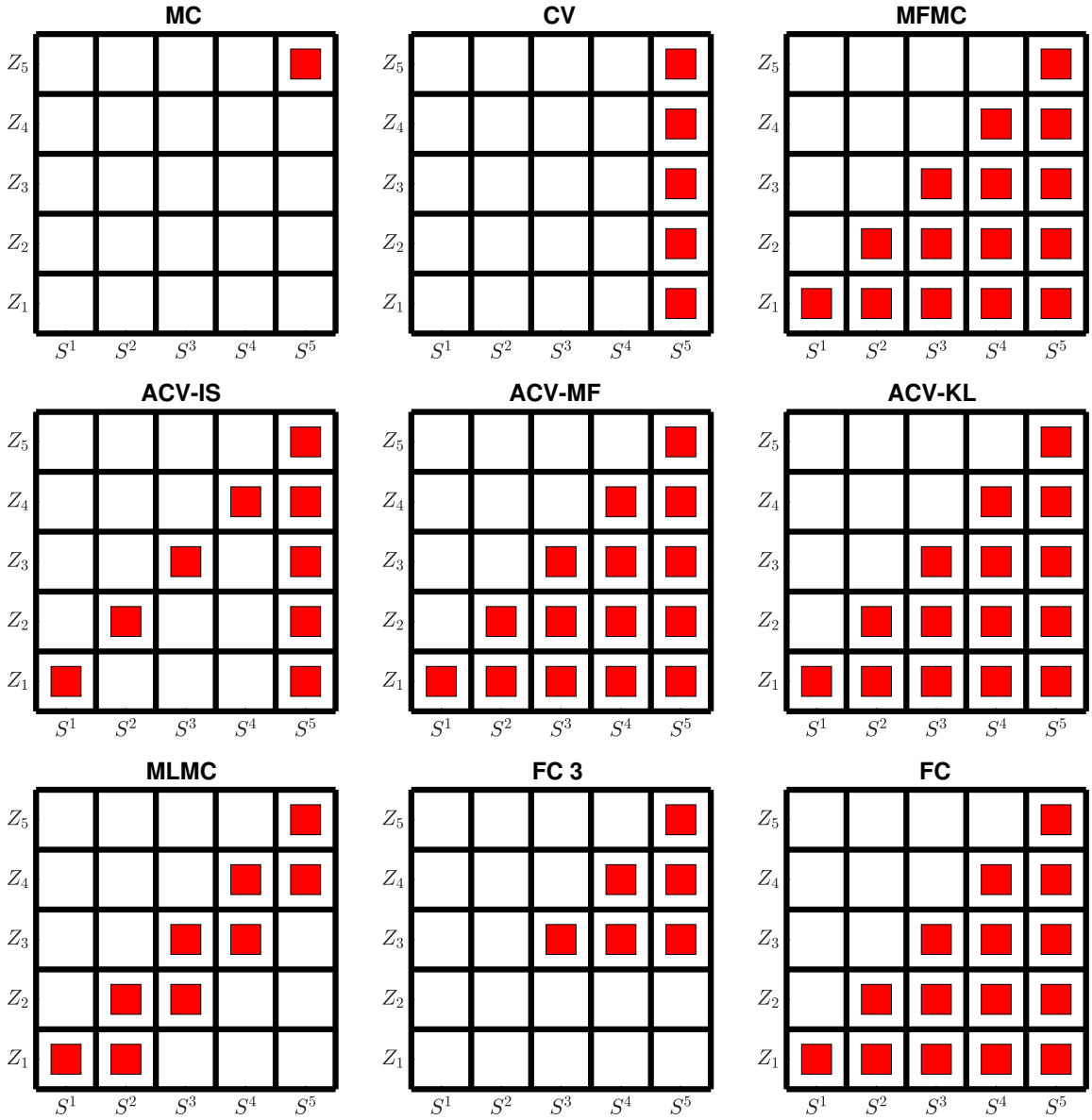

Figure 4.1: Model groups of different linear unbiased estimators $\widehat{\mu}_5$ with models $Z_1, \ldots, Z_5$. A model group S^k uses the model Z_ℓ if the respective square is red, otherwise it is white. The model groups S^k changes from estimator to estimator. We follow the convention that $Z_5 \in S^5$ for all estimators.

Condition	MC	CV	MFMC	ACV–IS	ACV–MF	ACV–KL	MLMC	FC
none	✓	✓	✗	✓	✗	✗	✗	✓
(4.40)	✓	✓	✓	✓	✓	✓	✗	✓
(4.41)	✓	✓	✗	✓	✓	✓	✗	✓

Table 4.1: Conditions under which the estimators are BLUE. "✓" means that the estimator is a BLUE and "✗" that the condition is not sufficient. A "✓" for the second row means that the estimator is unconditionally the BLUE. For $L = 1$ all estimators are equal to the MC estimator, thus all are the BLUE.

4.5 Numerical experiments

Monomial example. We define the QoI as monomials similar to [62, Section 2.5] for $L = 5$ as

$$Z_\ell(\omega) := \omega^\ell \qquad \text{for all } \ell \in \{1, \dots, 5\} \tag{4.42}$$

where the random source is uniformly distributed $\omega \sim U([0,1])$. We explicitly compute the mean, variance and covariance

$$\begin{aligned}
\mu_\ell &= \int_0^1 x^\ell \, dx = \frac{1}{\ell+1}, \\
\sigma_\ell^2 &= \frac{1}{2\ell+1} - \frac{1}{(\ell+1)^2}, \\
C_{\ell,j} &= \int_0^1 \left(x^\ell - \frac{1}{\ell+1}\right)\left(x^j - \frac{1}{j+1}\right) dx = \frac{1}{\ell+j+1} - \frac{1}{(\ell+1)(j+1)}.
\end{aligned}$$

and summarize them in Table 4.2. We derived in Section 4.3 that the BLUE uses the best approximation of Z_L in the space $\operatorname{span}(Z_1 - \mu_1, \dots, Z_{L-1} - \mu_{L-1})$ given enough low fidelity samples such that we can neglect the error in the estimation of $\mu_1, \dots, \mu_{L-1}$. Since the QoIs are monomials, the problem is equivalent to the approximation of a polynomial of higher order with polynomials of lower order on the interval $[0,1]$. The constant polynomial v is not used for the approximation since it would already approximate the mean, that is, the mean $v^* = \int_0^1 x^L dx = \mu_L$ is the unique minimizer of

$$\min_{v \in \mathbb{R}} \|Z_L - v\|^2_{L^2([0,1])} = \int_0^1 (x^L - v)^2 dx.$$

Our goal is to verify the lower variance bound $\mathbb{V}_I^{\min}$, which is plotted in Figure 4.2. We conclude that the highest variance reduction is achieved if we choose I to be as large as possible $I = \{1, \dots, L-1\}$ and that the inequality in Lemma 3.21 is true

$$\mathbb{V}_I^{\min} \leq \mathbb{V}_J^{\min} \qquad \text{for all } J \subseteq I.$$

We now want to estimate μ_L using the following number of evaluations of models Z_ℓ for different $N \in \mathbb{N}$

$$\begin{aligned}
n_\ell &:= 2^N 3^{L-\ell} \qquad \text{for all } \ell \in \{1, \dots, L-1\}, \\
n_L &:= 1.
\end{aligned}$$

	μ	σ^2
Z_1	0.50	0.08
Z_2	0.33	0.09
Z_3	0.25	0.08
Z_4	0.20	0.07
Z_5	0.17	0.06

Model P	Z_1	Z_2	Z_3	Z_4	Z_5
Z_1	1	0.97	0.92	0.87	0.82
Z_2	sym	1	0.99	0.96	0.93
Z_3	sym	sym	1	0.99	0.97
Z_4	sym	sym	sym	1	0.99
Z_5	sym	sym	sym	sym	1

Table 4.2: Monomial example: Mean, variance and Pearson correlation coefficient matrix P of the QoI defined in (4.42). The entry "sym" means symmetric and the value can be deduced from the relationship $\rho_{ij} = \rho_{ji}$.

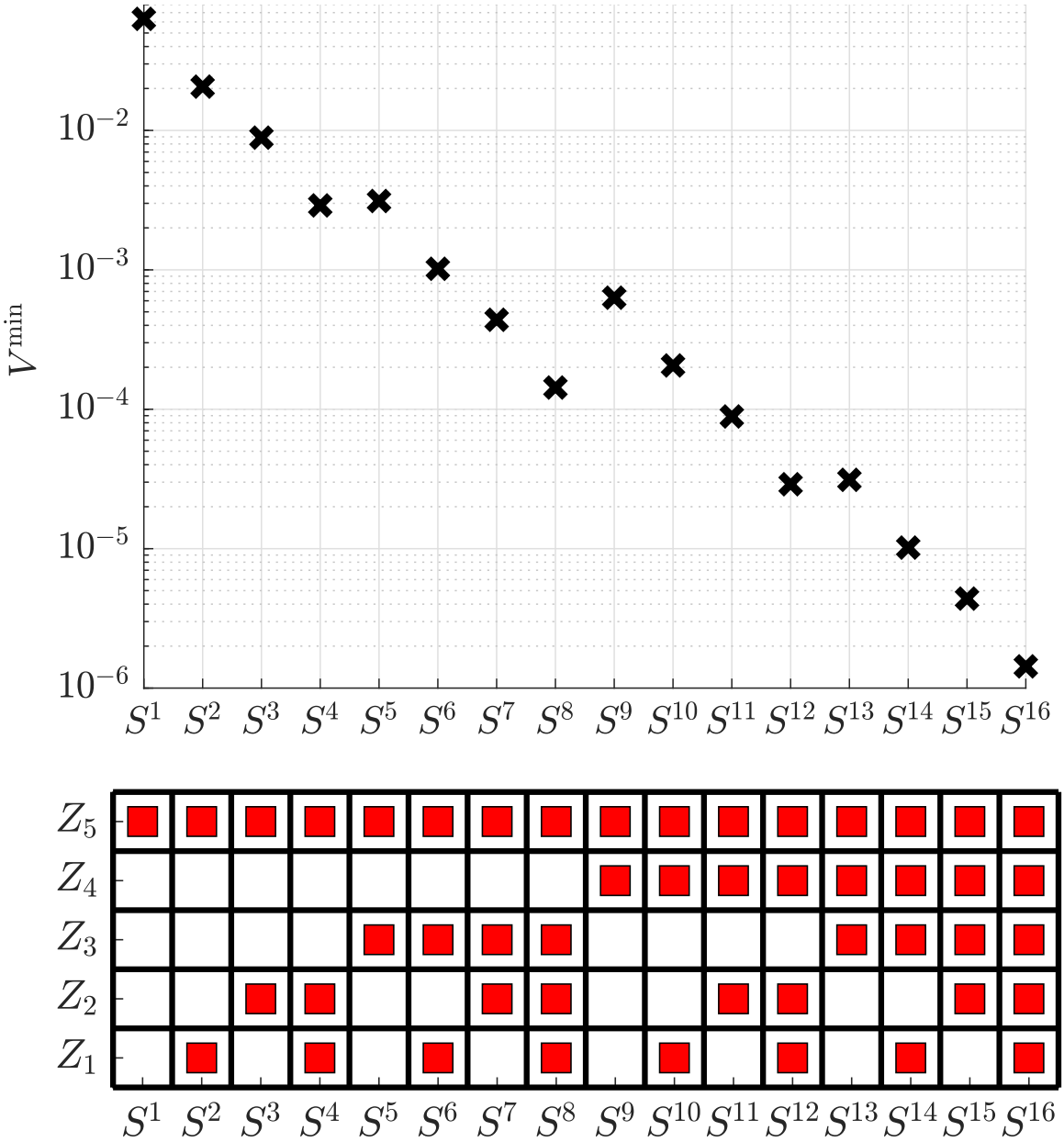

Figure 4.2: Monomial example: Different values of $\mathbb{V}_I^{\min}$ marked as "x" for different model groups $S = I \cup \{L\}$ that contain the high fidelity model Z_L. A red square means that the model is used whereas no square means that the model is not used.

We then choose sufficiently large N to simulate $m_\ell \to +\infty$ for $\ell \in \{1, \dots, L-1\}$, which corresponds to inexpensive low fidelity models. The number of evaluations of the high fidelity model is fixed to one, that is, $n_L = 1$.

We distribute the number of evaluations n_ℓ across the levels such that we have as much evaluations on the finer models as possible. As an example, for the MLMC estimator we use $m_L = n_L$ samples for $S^L = \{L-1, L\}$. There are now $m_{L-1} := n_{L-1} - n_L$ samples available for the evaluation of $S^{L-1} = \{L-2, L-1\}$. We continue with this scheme for all levels and apply this for every estimator. For the MFMC estimator the model group $S^L = \{1, \dots, L\}$ is evaluated only a single time since $n_L = 1$. This costs an evaluation of $Z_1, \dots, Z_{L-1}$, whereas the MLMC estimator only has an additional evaluation of Z_{L-1}. We do not use any model selection or optimize the estimator w.r.t. the overall cost. We

only optimize the coefficients β for the BLUE FC, CV, MFMC, ACV–IS, ACV–MF and ACV–KL estimator. We further optimize over the integer valued parameters K and N of the ACV–KL estimator.

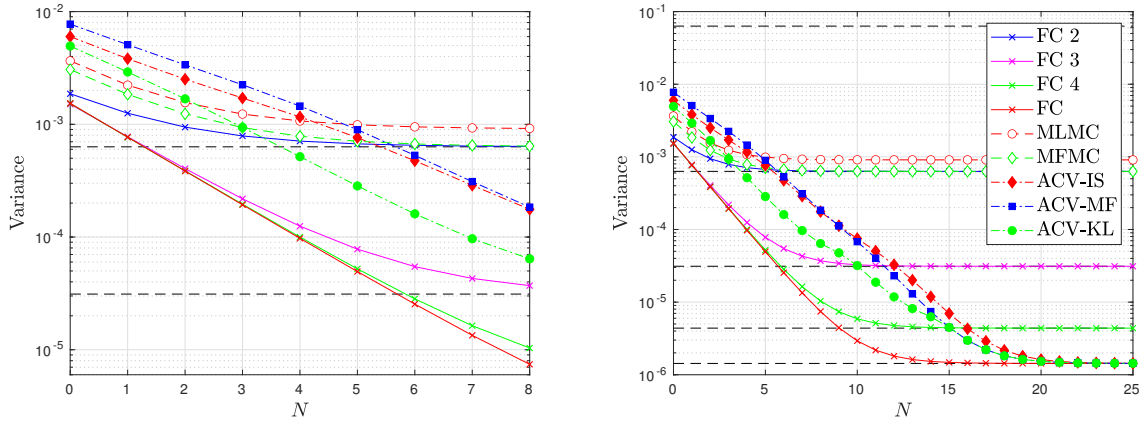

Figure 4.3: Monomial example: Variance of different estimators for different N. The left plot is a zoom in with $N \in \{0, \dots, 8\}$ of the right image with $N \in \{0, \dots, 25\}$. The dashed reference lines are $\mathbb{V}_{\emptyset}^{\min}, \mathbb{V}_{\{4\}}^{\min}, \mathbb{V}_{\{3,4\}}^{\min}, \dots, \mathbb{V}_{\{1,\dots,4\}}^{\min}$ with decreasing value. The MC estimator is not drawn, since its variance is equal to $\mathbb{V}_{\emptyset}^{\min}$. The CV estimators with control variates in I have variance equal to $\mathbb{V}_{I}^{\min}$.

We present the computed variance of different estimators in Figure 4.3. We conclude that the FC estimators have a smaller variance compared to the other estimators for small N. The FC estimator uses a sample allocation equal to the MFMC, ACV–MF and ACV–KL estimator and thus neither of these three estimator is in general the BLUE. We further conclude that the variance of the estimators is larger than the predicted minimal variance $\mathbb{V}^{\min}[\widehat{\mu}_L]$, even if we increase the number of low fidelity samples to infinity. The variance of the FC κ estimators converges to the predicted bound $\mathbb{V}_{\{L-\kappa+1,\dots,L\}}^{\min}$, in particular the FC estimator that couples all models is closest to the bound $\mathbb{V}_{\{1,\dots,L-1\}}^{\min}$. We remark that the variance of the MLMC estimator does not reach the respective bound $\mathbb{V}_{\{L-1\}}^{\min}$ and stops shortly before it.

We get a better explanation for this if we look at the coefficients in front of every sample group for every estimator for $N = 8$ in Figure 4.4. We conclude that the coefficient β^5 of S^5 of the FC estimator is very close to the coefficient of the CV estimator, which fits the theory. We further observe that the coefficients for MFMC, ACV–MF and ACV–KL have the same sign for all models Z_ℓ except one. This is in contrast to the FC estimators that allow models with multiple negative or positive coefficients, which leads to a checkerboard pattern for this example. We further remark that in the MFMC estimator the coefficients of the models in $S^5 = \{1, \dots 5\}$ are very close to zero. The result is that the models $Z_1, \dots, Z_{L-2}$ are not used for further variance reduction and thus the MFMC estimator is not able to achieve a variance smaller than $\mathbb{V}_{\{4\}}^{\min}$.

Noisy monomial example. The previous example showed the basic properties of the BLUE by optimally combining the models. A linear combination of the low fidelity models ensures that the residual has a small variance such that it is easy to estimate. We now give an example where the result is more pronounced by adding noise to the

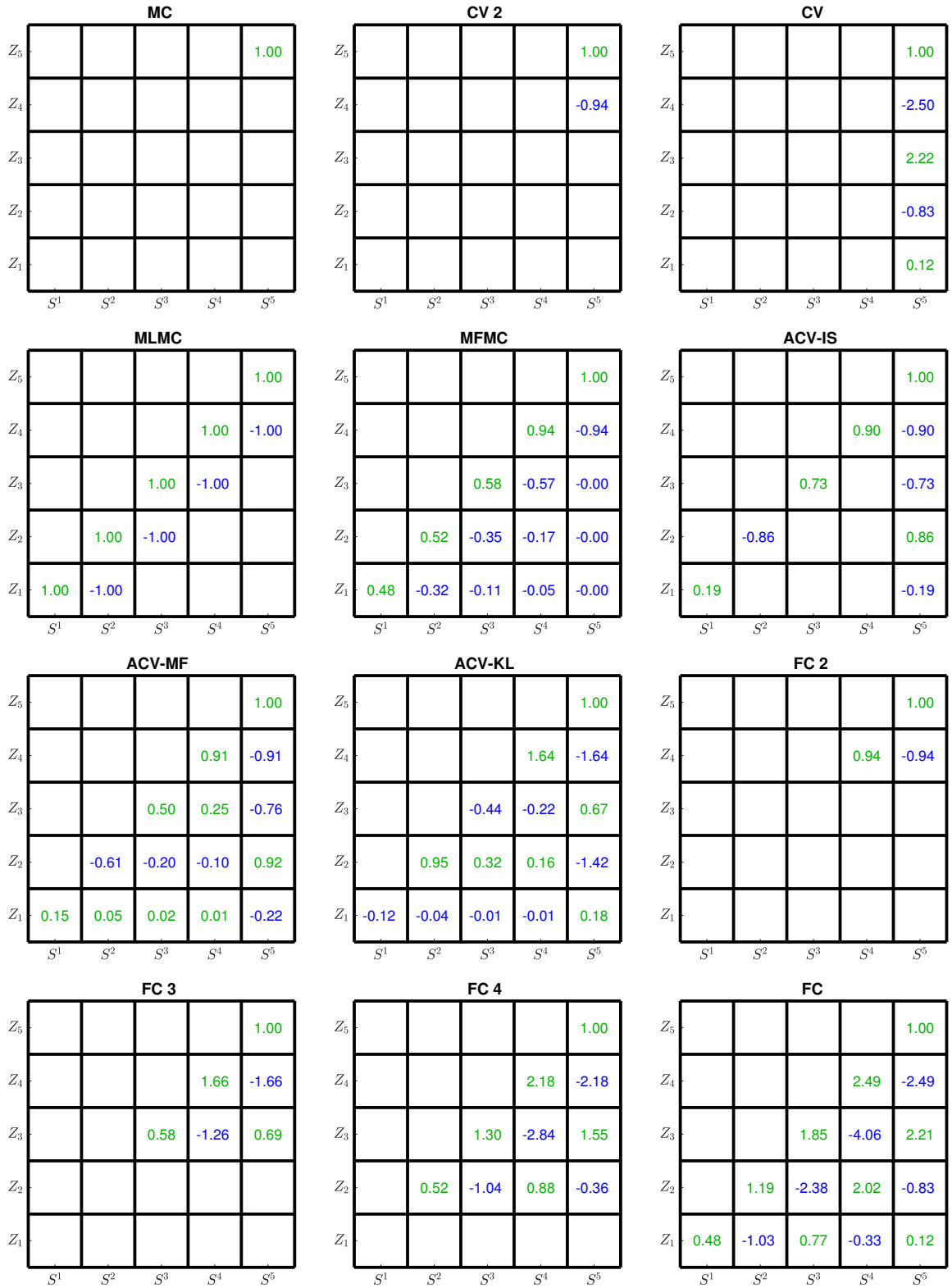

Figure 4.4: Monomial example: Coefficients of the linear combination for different linear unbiased estimators for $N = 8$. An empty box means that a model is not used for the respective model group. Each row has to sum up to 0 except for the top row which sums up to 1. Here we do not account for rounding errors and the CV estimator has the bias property only for the top row, since this estimator assumes that $\mu_1, \ldots, \mu_{L-1}$ is known. The estimator CV 2 uses a single control variate Z_4 and couples the two models Z_4 and Z_5. Green coefficients are positive and blue coefficients are negative.

monomials

$$\begin{aligned} Z_1(\omega) &:= 1 + \xi, \\ Z_\ell(\omega) &:= \omega^{\ell-1} + \xi \qquad \text{for all } \ell \in \{2, \dots, 5\}, \\ Z_6 &:= \omega^5. \end{aligned} \tag{4.43}$$

We again assume that $\omega \sim U([0,1])$ and that the noise $\xi \sim N(0,4)$ is independent of ω. The QoI Z_1 is up to a constant equal to the noise, $Z_2, \dots, Z_5$ have added noise and Z_6 has no noise and is equal to the high fidelity model for the monomial example without noise. The mean, variance and covariance can easily be computed using the independence of ω and ξ and we summarize these values in Table 4.3. We plot the different model groups and $\mathbb{V}_I^{\min}$ for different sets $I \subseteq \{1, \dots, L-1\}$ in Figure 4.5. We use the same sample allocation strategy as for the monomial example and show the variance of different estimators in Figure 4.6. Since the BLUE computes the best linear combination, we conclude that by defining $Z'_\ell := Z_\ell - Z_1$ the FC estimator removes the noise from the control variates $Z_2, \dots, Z_5$, therefore its performance should not change significantly compared to the monomial example without noise.
However, the other estimators, especially the MLMC and MFMC estimator, will have a large variance since the correlation $\rho_{L,L-1}$ is not at all close to one. Here the fixed a-priori ordering of these estimator leads to a large variance. The MLMC estimator has a far larger variance than even the MC estimator since $Z_L - Z_{L-1}$ is not small. This is typically detected by a model selection step, which could then use the MC estimator. Furthermore, every estimator that does not use the model Z_1 like the MC estimator or the FC κ estimators for $\kappa < 6$ are not able to significantly reduce the variance.
For a small number of low fidelity samples N the respective BLUE FC estimator has a much smaller variance than the other estimators.

Remark 4.31 (BLUE is robust)**.** We emphasize the robustness of the BLUE w.r.t. prior assumptions. The MLMC estimator makes the a-priori choice for the models and coefficients based on the idea that $Z_L - Z_{L-1}$ is small. If this is not satisfied, then the MLMC estimator may not even achieve a variance reduction. On the other hand the MLMC estimator does not require us to tune or compute any coefficients β, which typically requires the knowledge of the covariance matrix. ◇

Let us finally look at the coefficients for the estimators in Figure 4.7. The MFMC estimator chooses the coefficients to be close to zero except for the coefficient in front of Z_L,

	μ	σ^2
Z_1	1	4
Z_2	0.50	4.08
Z_3	0.33	4.09
Z_4	0.25	4.08
Z_5	0.20	4.07
Z_6	0.17	0.06

Model P	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
Z_1	1	0.99	0.99	0.99	0.99	0
Z_2	sym	1	1.00	1.00	1.00	0.12
Z_3	sym	sym	1	1.00	1.00	0.14
Z_4	sym	sym	sym	1	1.00	0.14
Z_5	sym	sym	sym	sym	1	0.13
Z_6	sym	sym	sym	sym	sym	1

Table 4.3: Noisy monomial example: Mean, variance and Pearson correlation coefficient matrix P of the QoI defined in (4.43). The entry "sym" means symmetric and its value can be deduced from the relationship $\rho_{ij} = \rho_{ji}$. Values with 1 are exactly 1 and values with 1.00 are rounded up from $0.99\dots$.

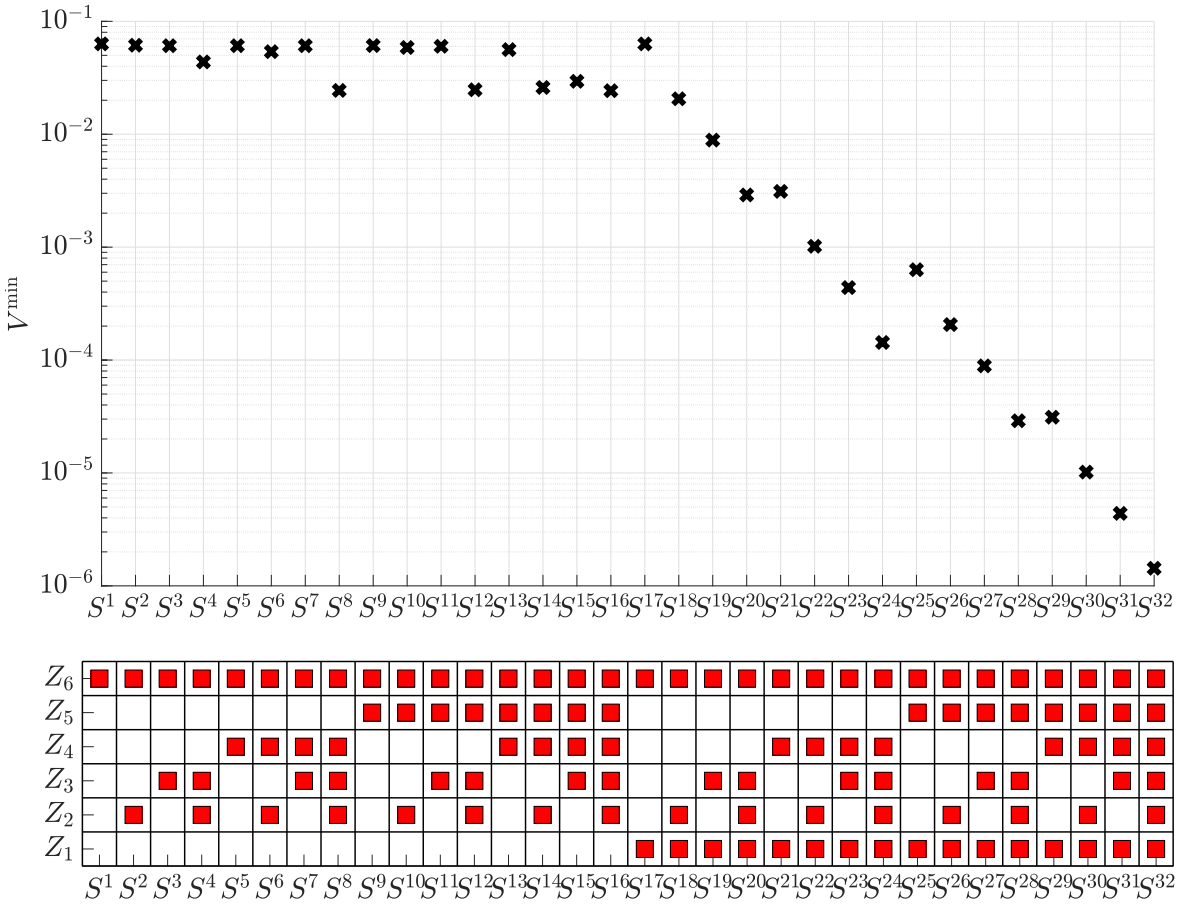

Figure 4.5: Noisy monomial example: Different values of $\mathbb{V}_I^{\min}$ marked as "x" for different model groups $S = I \cup \{L\}$ that contain the high fidelity model Z_L. A red square means that the model is used and an empty square that the model is not used.

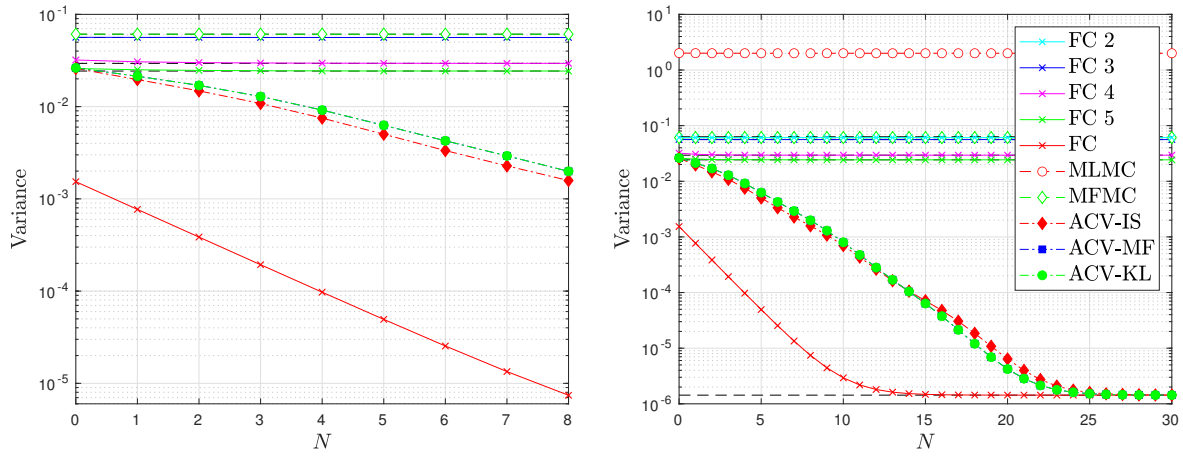

Figure 4.6: Noisy monomial example: Variance of different estimators for different N. The left plot is a zoom in with $N \in \{0, \ldots, 8\}$ of the right image with $N \in \{0, \ldots, 30\}$. The dashed reference lines are $\mathbb{V}_{\emptyset}^{\min}, \mathbb{V}_{\{5\}}^{\min}, \mathbb{V}_{\{4,5\}}^{\min}, \ldots, \mathbb{V}_{\{1,\ldots,5\}}^{\min}$ with decreasing value. The MC estimator is not drawn, since its variance is equal to $\mathbb{V}_{\emptyset}^{\min}$. The CV estimators with control variates in I have variance equal to $\mathbb{V}_I^{\min}$.

which does not reduce the variance in any significant way. If we add up the last column of the FC estimator w.r.t. the noise ξ we have

$$0\xi - 2.49\xi + 2.21\xi - 0.83\xi + 0.12\xi + 0.99\xi \approx 0,$$

which shows that the FC estimator removes the noise. We obtain a similar result for every column of the FC estimator. We further remark that the values in this matrix with rows $Z_2, \ldots, Z_6$ and columns $S^2, \ldots, S^6$ are almost identical to the values of the matrix of the FC estimator for the monomial example without noise in Figure 4.7. This confirms that the BLUE removes the noise from every sample group. The model Z_1 in S^1 of the FC estimator has a coefficient very close to zero, which means that this sample group is essentially not used. This is not surprising, since Z_1 and Z_6 are independent and we can always subtract a multiple of Z_1 and thus ξ in the linear combination $(\beta^k)^T Z_{S^k}$ due to the model group structure $S^\ell = \{1, \ldots, \ell\}$.

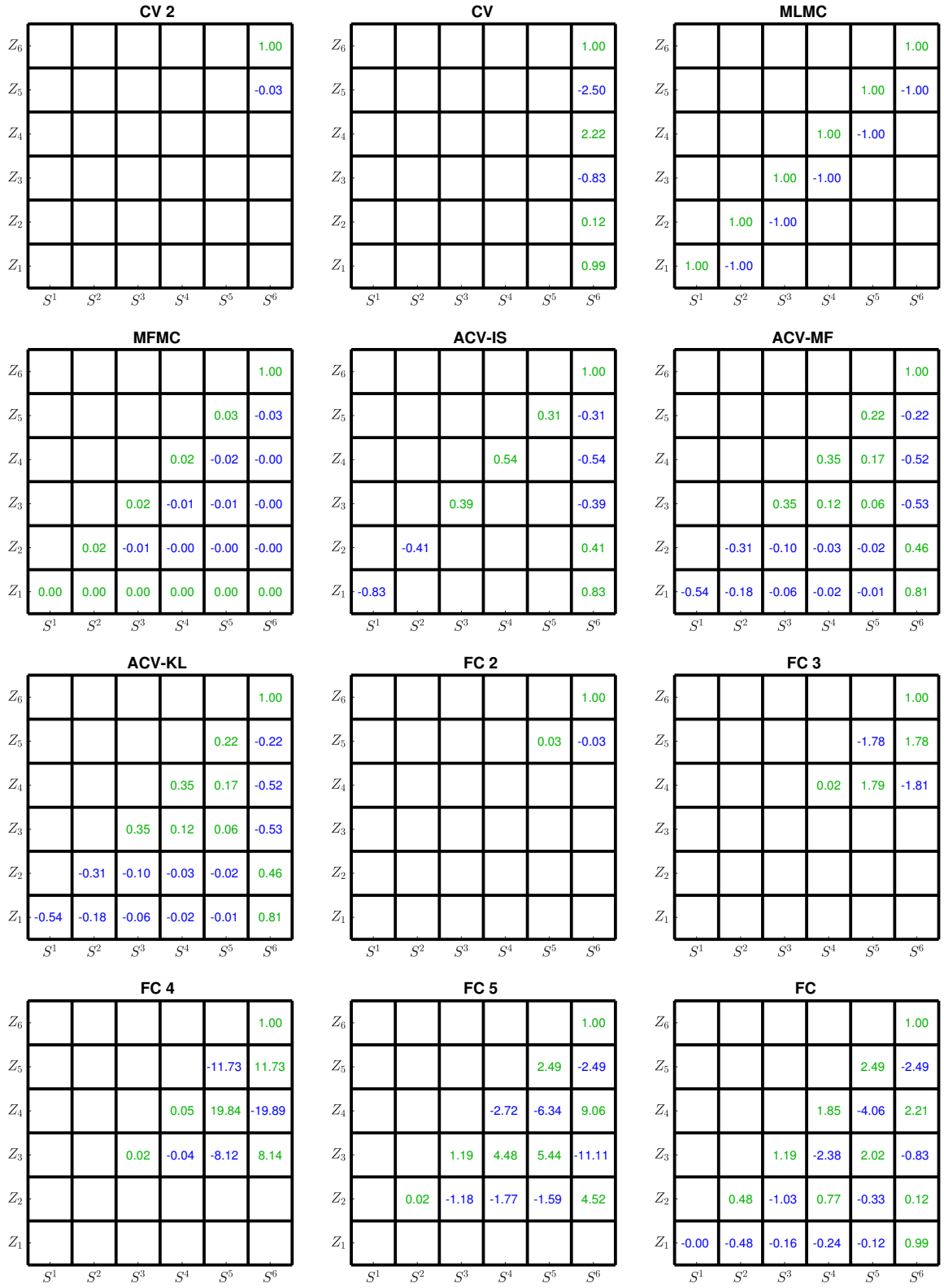

Figure 4.7: Noisy monomial example: Coefficients of the linear combination for different linear unbiased estimators for $N = 8$. An empty box means that a model is not used for the respective model group. Each row has to sum up to 0 except for the top row, which sums up to 1. Here we do not account for rounding errors and the CV estimator has the bias property only for the top row, since this estimator assumes that $\mu_1, \ldots, \mu_{L-1}$ is known. Green coefficients are positive and blue coefficients are negative.

Chapter 5

Sample allocation optimal BLUE

The condition that an estimator is the BLUE is not a sufficient for being an inexpensive estimator. Chapter 4 showed that the MC estimator is the BLUE in contrast to the MLMC estimator which is in general not a BLUE. Nevertheless, MLMC has a better asymptotic complexity than MC if we compare Theorem 3.11 with Theorem 3.49. The reason is that both estimators use different samples such that the achievable variance is smaller for the MLMC estimator. Therefore we have to equip the BLUE with a method to choose the linear regression problem which we call the sample allocation problem. In this chapter we proceed as follows:

- We define the sample allocation problem in Section 5.1. We further prove the existence of a solution which defines a sample allocation optimal BLUE and verify that this estimator is optimal in the class of linear unbiased estimators. We further extend the sample allocation problem to include a coupling number κ such that only model groups are evaluated with less or equal to κ models.
- We relax the sample allocation problem to obtain a tractable solution and discuss why this relaxation is sensible. The relaxed problem in Section 5.2 has a solution that is in general not unique. We further show that optimal solutions are sparse in the sense that at most L of the $K = 2^L - 1$ model groups are used. This property is crucial for a practical implementation due to the necessity for rounding the number of samples to the next integer.
- Section 5.3 provides analogous but slightly stronger results compared to Section 5.2. The main idea is to view the sample allocation problem from a different angle and first optimize the sample allocation and afterwards the coefficient of the linear unbiased estimator. These results simplify the asymptotic analysis in Chapter 6.
- In Section 5.4 we prove that every minimizer of the sample allocation problem lies in the convex hull of a finite number of minimizers where each of these uses at most L model groups.

The sections of this chapter in terms of their minimization problems are given in Figure 5.1. This chapter contains results from [125, 126] and Section 5.2 together with Section 5.4 extend some of these results.

5.1 Ideal sample allocation optimal BLUE

We know from Chapter 4 that the variance of the BLUE with coefficients β^k in (4.15) and number of samples m is

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}\big] = \sum_{k \in U_S} \frac{(\beta^k)^T C^k \beta^k}{m_k} = \alpha^T \Psi^{-1} \alpha = \alpha^T \left(\sum_{k \in U_S} m_k P^k (C^k)^{-1} R^k \right)^{-1} \alpha.$$

In the following we ease the notation and make the dependence on the number of samples m of the system matrix $\Psi(m)$ and the estimator $\widehat{\mu}_\alpha^{\mathrm{B}}(m)$ explicit. The goal is to choose m

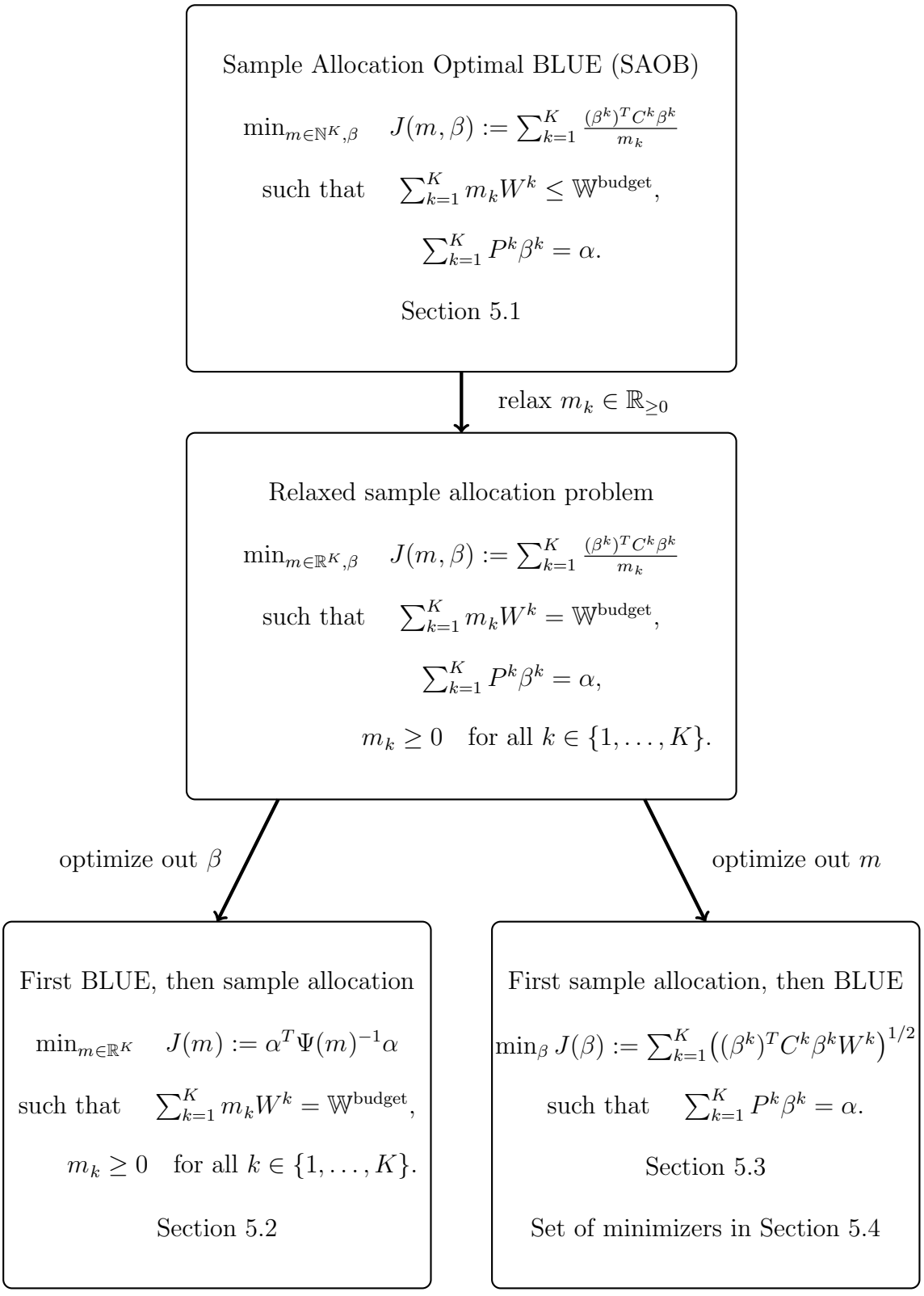

Figure 5.1: Sections of Chapter 5 described in terms of the respective minimization problem that we examine. Here $\mathbb{W}^{\text{budget}} > 0$ denotes a computational bugdet and W^k the cost to compute a single evaluation of all models in the model groups S^k. There are two methods to derive an approximation of the SAOB: The common step is to first allow non–integer samples to obtain a tractable optimization problem. Then we use explicit expression for the optimal β or m to obtain two different minimization problems which both lead to the SAOB. These are examined in Section 5.2 and Section 5.3.

such that the variance is minimized and we do not exceed a known computational budget similar to the MLMC, MFMC and ACV estimators in Chapter 3. This cost constraint is necessary since otherwise the variance can be made as small as desired by choosing m sufficiently large. We define the cost of a model group as the sum of the costs of every contained model

$$W^k := \mathbb{W}[Z_{S^k}] = \sum_{\ell \in S^k} \mathbb{W}[Z_\ell] = \sum_{\ell \in S^k} w_\ell \qquad \text{for all } k \in \{1, \dots, K\}.$$

We introduce a fixed budget $\mathbb{W}^{\text{budget}} > 0$ and write down the sample allocation problem

$$\begin{aligned} \min_{m_1,\dots,m_K \in \mathbb{N}_0} \quad & \mathbb{V}\big[\widehat{\mu}_\alpha^{\text{B}}(m)\big] = \alpha^T \Psi(m)^{-1} \alpha \\ \text{such that} \quad & \mathbb{W}\big[\widehat{\mu}_\alpha^{\text{B}}(m)\big] = \sum_{k \in \{1,\dots,K\}} m_k W^k \leq \mathbb{W}^{\text{budget}}. \end{aligned} \tag{5.1}$$

We define $\mathbb{V}\big[\widehat{\mu}_\alpha^{\text{B}}(m)\big] := +\infty$ if m is such that the BLUE $\widehat{\mu}_\alpha^{\text{B}}(m)$ is not well defined. The BLUE is not necessarily unique, therefore we are only able to prove existence of a solution of (5.1).

Theorem 5.1 (Existence of sample allocation optimal BLUE)**.** Let $\mathbb{W}^{\text{budget}}$ be such that we are able to evaluate the required models at least once

$$\sum_{\ell \in U_\alpha} w_\ell \leq \mathbb{W}^{\text{budget}}. \tag{5.2}$$

Then there exists a minimizer m^* of (5.1) and every corresponding BLUE is a *sample allocation optimal BLUE (SAOB)*

$$\widehat{\mu}_\alpha^{\text{SAOB}} := \widehat{\mu}_\alpha^{\text{B}}(m^*).$$

Proof. The sample allocation problem (5.1) optimizes over a finite set of $m \in \mathbb{N}_0^K$ since the cost of every model $w_\ell > 0$ is positive. We thus enumerate all values and choose m^* such that the variance is minimized. Assumption (5.2) now ensures that the MC estimator with a single sample defines a feasible BLUE

$$\widehat{\mu}_\alpha^{\text{MC}} := \alpha^T Z_{U_\alpha}^1.$$

Hence we conclude that m^* is feasible and thus $\mathbb{V}\big[\widehat{\mu}_\alpha^{\text{B}}(m^*)\big] = \mathbb{V}\big[\widehat{\mu}_\alpha^{\text{SAOB}}\big] < +\infty$. □

The SAOB is optimal in the class of linear unbiased estimators. No other linear unbiased estimator is able to achieve a smaller variance satisfying the cost constraint.

Theorem 5.2 (Optimality of SAOB)**.** Let $\widehat{\mu}_\alpha$ be a linear unbiased estimator that uses m_k i.i.d. samples of S^k for $k \in \{1, \dots, K\}$, where the samples are also independent across model groups. Then every SAOB with $\mathbb{W}^{\text{budget}} = \mathbb{W}[\widehat{\mu}_\alpha]$ has smaller or equal variance

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\text{SAOB}}\big] \leq \mathbb{V}[\widehat{\mu}_\alpha].$$

Proof. The BLUE $\widehat{\mu}_\alpha^{\text{B}}(m)$ with the same sample allocation m as $\widehat{\mu}_\alpha$ has smaller or equal variance

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\text{B}}(m)\big] \leq \mathbb{V}[\widehat{\mu}_\alpha].$$

The Assumption (5.2) is satisfied since $\widehat{\mu}_\alpha^{\text{B}}$ is a feasible estimator in (5.1) and thus a SAOB exists. However, the SAOB is the minimizer of this problem and thus we conclude the theorem

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\text{SAOB}}\big] = \mathbb{V}\big[\widehat{\mu}_\alpha^{\text{B}}(m^*)\big] \leq \mathbb{V}\big[\widehat{\mu}_\alpha^{\text{B}}(m)\big] \leq \mathbb{V}[\widehat{\mu}_\alpha].$$ □

The two previous theorems allow the usage of every model group S^k and similar results are achievable if we allow only the use of specific model groups. We are particularly interested in evaluating or coupling at most κ models in a single model group, that is $m_k = 0$ if $|S^k| > \kappa$. The *SAOB with coupling* $\kappa \in \mathbb{N}$ (SAOB κ) is a BLUE with a sample allocation that solves

$$\begin{aligned} \min_{m_1,\dots,m_K \in \mathbb{N}_0} \quad & \mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}(m)\big] = \alpha^T \Psi(m)^{-1} \alpha \\ \text{such that} \quad & \sum_{k \in \{1,\dots,K\}} m_k W^k \leq \mathbb{W}^{\mathrm{budget}}, \\ & m_k = 0 \qquad \text{for } k \in \{1,\dots,K\} \text{ with } |S^k| > \kappa. \end{aligned}$$

We denote this estimator with $\widehat{\mu}^{\mathrm{SAOB}\,\kappa}$ and remark that its existence and optimality follows analogously to SAOB. For completeness we write down the optimality result.

Corollary 5.3 (Optimality of SAOB κ)**.** Let $\widehat{\mu}_\alpha$ be a linear unbiased estimator that uses m_k i.i.d. samples of S^k for $k \in \{1,\dots,K\}$ with $|S^k| \leq \kappa$, where the samples are also independent across model groups. Then every SAOB κ with $\mathbb{W}^{\mathrm{budget}} = \mathbb{W}[\widehat{\mu}_\alpha]$ has smaller or equal variance

$$\mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa}\big] \leq \mathbb{V}[\widehat{\mu}_\alpha]. \qquad \diamond$$

The coupling parameter κ of SAOB κ is important for the asymptotic analysis in Chapter 6. We further remark that κ also determines the lower variance bound of SAOB κ, since we are allowed to couple at most $\kappa - 1$ other models with the high fidelity model

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\mathrm{SAOB}\,\kappa}] = \min_{\substack{I \subseteq \{1,\dots,L\}, \\ |I| = \kappa - 1}} \mathbb{V}_I^{\min}.$$

It is possible to weaken the assumptions of i.i.d. samples for every model group and across different model groups. The BLUE is the solution of a generalized linear regression problem, where a different sample structure leads to a different noise assumption, hence a BLUE still exists. We are further able to generalize the cost constraint by simply bounding the cost of the respective BLUE

$$\begin{aligned} \min_{\widehat{\mu}_\alpha^{\mathrm{B}}} \quad & \mathbb{V}\big[\widehat{\mu}_\alpha^{\mathrm{B}}(m)\big] \\ \text{such that} \quad & \mathbb{W}\big[\widehat{\mu}_\alpha^{\mathrm{B}}(m)\big] \leq \mathbb{W}^{\mathrm{budget}}. \end{aligned}$$

This problem typically has a feasible point if $\mathbb{W}^{\mathrm{budget}}$ is large enough. Since the number of samples are integer, the cost constraints often ensures that the number of evaluations and thus the number of feasible BLUEs is bounded. We further have to demand that the estimator $\widehat{\mu}_\alpha^{\mathrm{B}}(m)$ only depends on m and no further other independent parameters that may modify the noise or covariance structure. Then the solution of this problem is again a SAOB that is now optimal in this defined class of linear unbiased estimators with costs bounded by $\mathbb{W}^{\mathrm{budget}}$.

The above generalization is interesting but in its generality not very useful. In practice, we have to compute an approximation to the minimization (5.1), which we will do by allowing the number of samples to be positive real numbers. If we additionally change the structure of the noise η for the regression problem in Definition 4.11 to include dependencies across model groups, then the dependence of the system matrix Ψ w.r.t. m is more complex. Furthermore, the covariance matrix C is often unknown in practice and introducing more coupling requires that we estimate more entries of the covariance matrix of the noise η.

5.2 First BLUE, then sample allocation

Relaxed formulation. We state reasons how and why we simplify the optimization problem (5.1). We relax the constraint that the number of samples m are integer to non–negative real numbers to obtain a problem that is numerically tractable

$$\begin{aligned} \min_{m_1,\dots,m_K\in\mathbb{R}} \quad & J(m) := \alpha^T \Psi(m)^{-1}\alpha \\ \text{such that} \quad & \sum_{k=1}^{K} m_k W^k = \mathbb{W}^{\text{budget}}, \\ & m_k \geq 0 \qquad \text{for all } k \in \{1,\dots,K\}. \end{aligned}$$

We have replaced the inequality for the cost constraint with equality which is always satisfied at every minimizer. This problem is still difficult to handle, since Ψ might not be well defined or not invertible, however the BLUE might still exist. We first show an example where this is not a problem.

Example 5.4 (Ψ not invertible)**.** Let $L := 2$, $w_1 := w_2 := 1$, $\mathbb{W}^{\text{budget}} := 3$, $\alpha := (0,1)^T$ and let the covariance matrix be the identity $C := I$. Then the MC estimator with three samples of Z_2 is the BLUE since samples of Z_1 never decrease the variance. This can formally be checked by looking at the lower variance bound $\mathbb{V}_{\emptyset}^{\min} = \mathbb{V}_{\{1\}}^{\min}$. Notice however that the matrix

$$\Psi = m_1 P^1 (C^1)^{-1} R^1 = 3\begin{pmatrix}0\\1\end{pmatrix}1\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 3\end{pmatrix}$$

is not invertible on $\mathbb{R}^2$ but on $V := \operatorname{span}(\alpha)$, where it matters, that is

$$\Psi : V \to V, \quad \Psi\alpha = 3\alpha.$$

We are thus able to compute the inverse of the linear operator Ψ on V and thus

$$\alpha^T \Psi^{-1} \alpha = \frac{1}{3}\|\alpha\|^2 = \frac{1}{3},$$

which is the variance of the MC estimator with three samples of Z_2. ◇

The second and more subtle problem is that the BLUE is not well defined if we do not evaluate all models that are required by the bias. However, since we demand $m \geq 0$ the number of evaluations of a model may tend to zero if the respective covariance matrices are close to non–invertibility. We demonstrate this in the next example.

Example 5.5 (C not invertible)**.** Let $L := 2$, $w_1 := w_2 := 1$, $\mathbb{W}^{\text{budget}} := 1$, $\alpha := (1,1)^T$ with covariance matrix $C := \begin{pmatrix}\varepsilon^2 & 0\\ 0 & 1\end{pmatrix}$. It is straightforward to see that we obtain the smallest variance if we evaluate the uncorrelated Z_1 and Z_2 separately. We compute the system matrix

$$\Psi(m) = m_1 P^1 (C^1)^{-1} R^1 + m_2 P^2 (C^2)^{-1} R^2 = \begin{pmatrix} m_1\varepsilon^{-2} & 0\\ 0 & m_2\end{pmatrix}.$$

This matrix is invertible for all $\varepsilon^2 > 0$ and the variance of the BLUE is

$$J(m) = \alpha^T \Psi(m)^{-1}\alpha = \begin{pmatrix}1\\1\end{pmatrix}^T \begin{pmatrix}\frac{1}{m_1}\varepsilon^2 & 0\\ 0 & \frac{1}{m_2}\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{m_1}\varepsilon^2 + \frac{1}{m_2}.$$

We minimize this expression w.r.t. the cost constraint and use the KKT conditions such that the optimal m_1 and m_2 satisfy

$$\frac{1}{m_1^2}\varepsilon^2 = \frac{1}{m_2^2},$$
$$m_1 + m_2 = \mathbb{W}^{\text{budget}}.$$

The optimal solution and the variance is then

$$m_1 = \frac{\varepsilon}{1+\varepsilon}, \quad m_2 = \frac{1}{1+\varepsilon}, \quad J(m) = (1+\varepsilon)\varepsilon + (1+\varepsilon).$$

The above expressions are well defined for all $\varepsilon \geq 0$ and in particular for $\varepsilon = 0$ for which C is not invertible. The latter is problematic, since then $m_1 = 0$ and thus we obtain an estimator with $U_\alpha = \{1,2\} \not\subseteq U_Z = \{2\}$. Therefore, the estimator $\widehat{\mu}_\alpha^{\text{B}}$ for $\varepsilon = 0$ with $m_1 = 0$ and $m_2 = 1$ is not a well defined linear unbiased estimator of $\alpha^T\mu$. ◇

Formulated differently, the set where the bias constraint can be satisfied is not closed

$$\left\{ m_1, \dots, m_K \in \mathbb{R}_{\geq 0} \,|\, \text{there exists } \beta^k \text{ with } \alpha = \sum_{\{k \,|\, m_k > 0\}} P^k \beta^k \right\}$$

and minimizing J leads to a point not in this set. We circumvent this difficulty if we assume that we pay the price of having at least a single evaluation of all models $S^L := \{1, \dots, L\}$.
Intuitively, it does not make sense to evaluate a model group S^k where C^k is not positive definite, since we know that the vectors $(Z_\ell - \mu_\ell)_{\ell \in S^k}$ are linearly dependent. Ignoring the mean μ for now, we are able to compute the value of any linear combination of $(Z_\ell)_{\ell \in S^k}$ without evaluating a particular model $j \in S^k$ and thus we do not have to pay the cost w_j. We formalize this intuition where we do not ignore the mean μ.

Lemma 5.6 (Model groups not used)**.** Let $\widehat{\mu}_\alpha$ be a linear unbiased estimator with $m_L \geq 1$ for the model group $S^L := \{1, \dots, L\}$ and C not positive definite. Then there exists an estimator $\widehat{\mu}'_\alpha$ with sample allocation m with equal variance $\mathbb{V}[\widehat{\mu}'_\alpha] = \mathbb{V}[\widehat{\mu}_\alpha]$ and smaller or equal cost $\mathbb{W}[\widehat{\mu}'_\alpha] \leq \mathbb{W}[\widehat{\mu}_\alpha]$ such that all model groups S^k where C^k is not positive definite are not used, or used only once for $k = L$

$$m_k = \begin{cases} 0, & \text{if } k \neq L, \\ 1, & \text{if } k = L. \end{cases}$$

Proof. Let $k \in \{1, \dots, K\} \setminus \{L\}$ such that C^k is not positive definite and assume that $\widehat{\mu}_\alpha$ uses this model group. Lemma 2.16 now shows that there exist coefficients $(a_\ell)_{\ell \in S^k} \neq 0$ such that almost surely

$$\sum_{\ell \in S^k} a_\ell Z_\ell = \sum_{\ell \in S^k} a_\ell \mu_\ell. \tag{5.3}$$

This allows us to represent an arbitrary model $j \in S^k$ with $a_j \neq 0$ in terms of the other QoIs and the mean values

$$Z_j = -\sum_{\ell \in S^k \setminus \{j\}} \frac{a_\ell}{a_j} Z_\ell + \mu_j + \sum_{\ell \in S^k \setminus \{j\}} \frac{a_\ell}{a_j} \mu_\ell = -\sum_{\ell \in S^k \setminus \{j\}} \frac{a_\ell}{a_j} Z_\ell + \sum_{\ell \in S^k} \frac{a_\ell}{a_j} \mu_\ell. \tag{5.4}$$

W.l.o.g. we assume that the linear unbiased estimator $\widehat{\mu}_\alpha$ uses equal coefficients for every sample in the model group S^k

$$\widehat{\mu}_\alpha = \sum_{n\in U_S\setminus\{k\}} \sum_{\ell\in S^n} \beta^n_{[\ell]} \left(\frac{1}{m_n} \sum_{i=1}^{m_n} Z_\ell^{i,n} \right) + \sum_{\ell\in S^k} \beta^k_{[\ell]} \left(\frac{1}{m_k} \sum_{i=1}^{m_k} Z_\ell^{i,k} \right)$$

and we replace Z_j in the last expression with (5.4)

$$\sum_{\ell\in S^k\setminus\{j\}} \beta^k_{[\ell]} Z_\ell^{i,k} + \beta^k_{[j]} Z_j^{i,k} = \sum_{\ell\in S^k\setminus\{j\}} \left(\beta^k_{[\ell]} - \frac{a_\ell}{a_j}\beta^k_{[j]} \right) Z_\ell^{i,k} + \beta^k_{[j]} \sum_{\ell\in S^k} \frac{a_\ell}{a_j}\mu_\ell.$$

This replacement does not change the variance and we define the estimator $\widehat{\mu}^*_\alpha$ to be

$$\widehat{\mu}^*_\alpha := \sum_{n\in U_S\setminus\{k\}} \sum_{\ell\in S^n} \beta^n_{[\ell]} \frac{1}{m_n} \sum_{i=1}^{m_n} Z_\ell^{i,n} + \sum_{\ell\in S^k\setminus\{j\}} \left(\beta^k_{[\ell]} - \frac{a_\ell}{a_j}\beta^k_{[j]} \right) \frac{1}{m_k} \sum_{i=1}^{m_k} Z_\ell^{i,k} + \beta^k_{[j]} \sum_{\ell\in S^k} \frac{a_\ell}{a_j}\mu_\ell.$$

The last part of this estimator is a linear combination of the mean values, which we know from (5.3) if we have a sample of $S^L = \{1,\dots,L\}$ available. Moreover, the estimator $\widehat{\mu}^*_\alpha$ is cheaper than $\widehat{\mu}_\alpha$ since it does not evaluate Z_j in the model group S^k

$$\mathbb{W}[\widehat{\mu}_\alpha] = \mathbb{W}[\widehat{\mu}^*_\alpha] - m_k \mathbb{W}\big[S^k\big] + m_k \mathbb{W}\big[S^k \setminus \{j\}\big] = \mathbb{W}[\widehat{\mu}^*_\alpha] - m_k w_j. \tag{5.5}$$

We now repeat the procedure outlined in this proof for $\widehat{\mu}^*_\alpha$ until we obtain an estimator $\widehat{\mu}'_\alpha$ such that $m_k = 0$ for all $k \in \{1,\dots,K\}\setminus\{L\}$ where C^k is not positive definite. We remark that this approach works even if we remove only some samples of a model group with non–positive definite covariance matrix while keeping the remaining ones. Since this proof is valid if S^L is evaluated at least once, we keep a single sample $m_L = 1$ for $\widehat{\mu}'_\alpha$. □

The above lemma is a justification to only look at model groups where the respective covariance matrix is positive definite. We view this as preliminary model selection where we are able to ignore some model groups without increasing the variance of the minimizer. We incorporate the evaluation of S^L in an alternative way into the cost function by adding the diagonal matrix δI with $\delta > 0$ such that $\Psi_\delta(m) := \Psi(m) + \delta I$ is invertible. We interpret δI as additional independent evaluations of $Z_1,\dots,Z_L$ in the following sense:

$$\delta I = \sum_{\ell=1}^{L} \delta\sigma_\ell^2 P^\ell (C^\ell)^{-1} R^\ell.$$

We used that $S^\ell = \{\ell\}$ and assumed $\sigma_\ell^2 > 0$ for $\ell \in \{1,\dots,L\}$, hence adding δI corresponds to $\delta\sigma_\ell^2$ additional evaluations of Z_ℓ. We then define the relaxed sample allocation problem

$$\begin{aligned} \min_{m_1,\dots,m_K\in\mathbb{R}} \quad & J_\delta(m) := \alpha^T \Psi_\delta(m)^{-1}\alpha \\ \text{such that} \quad & \sum_{k=1}^{K} m_k W^k = \mathbb{W}^{\text{budget}}, \\ & m_k \geq 0 \qquad \text{for all } k \in \{1,\dots,K\}. \end{aligned} \tag{5.6}$$

The matrix Ψ_δ and its inverse is well defined if C is positive definite or if we restrict ourselves to model groups where C^k is positive definite. In the latter setting we however, may have to pay the price of an additional evaluation of all models.

Remark 5.7 (Other relaxations)**.** We are able to replace δI by δC^{-1} to obtain a δ–th of an evaluation of $\{1,\dots,L\}$ if C is positive definite. Another alternative is to ensure that all models are evaluated at least δ-times

$$\sum_{\{k\in\{1,\dots,K\}\,|\,\ell\in S^k\}} m_k \geq \delta \quad \text{for all } \ell \in \{1,\dots,L\}.$$

Both methods complicate the analysis and our numerical results in Chapter 6 suggest that setting $\delta = 0$ is possible. Furthermore, in Section 5.3 we apply a different approach which allows us to remove the matrix δI altogether. ◇

We derive a lower and upper bound on the error of the variance for replacing Ψ with Ψ_δ. Since δI can be viewed as additional samples, we expect that the error gets smaller and smaller if we use more and more samples.

Lemma 5.8 (Variance with added δI)**.** For positive definite Ψ and $\delta \geq 0$ the change in the variance by adding δI satisfies

$$\frac{1}{1+\delta\lambda_{\max}(\Psi^{-1})}\alpha^T\Psi^{-1}\alpha \leq \alpha^T\Psi_\delta^{-1}\alpha \leq \frac{1}{1+\delta\lambda_{\min}(\Psi^{-1})}\alpha^T\Psi^{-1}\alpha, \tag{5.7}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest respectively largest eigenvalue of Ψ. In particular, if we increase the number of samples to infinity the relative error converges to zero in the following sense

$$\frac{\alpha^T\Psi(sm)^{-1}\alpha - \alpha^T\Psi_\delta(sm)^{-1}\alpha}{\alpha^T\Psi(sm)^{-1}\alpha} \leq \frac{\delta\lambda_{\max}(\Psi^{-1})}{s+\delta\lambda_{\max}(\Psi^{-1})} \to 0 \quad \text{for } s\to+\infty. \tag{5.8}$$

Proof. We extract the matrix $\Psi^{-1/2}$ and use the bound of the largest eigenvalue

$$\begin{aligned}\alpha^T{\Psi_\delta}^{-1}\alpha = \alpha^T(\Psi+\delta I)^{-1}\alpha &= \alpha^T\Psi^{-1/2}(I+\delta\Psi^{-1})^{-1}\Psi^{-1/2}\alpha \\ &\leq \lambda_{\max}((I+\delta\Psi^{-1})^{-1})\alpha^T\Psi^{-1}\alpha.\end{aligned}$$

For positive definite matrices the largest eigenvalue of its inverse is the inverse of the smallest eigenvalue

$$\lambda_{\max}((I+\delta\Psi^{-1})^{-1}) = \frac{1}{\lambda_{\min}(I+\delta\Psi^{-1})} = \frac{1}{1+\delta\lambda_{\min}(\Psi^{-1})},$$

which shows the upper bound in (5.7). For the lower bound we use the inequality

$$\alpha^T\Psi^{-1/2}(I+\delta\Psi^{-1})^{-1}\Psi^{-1/2}\alpha \geq \lambda_{\min}((I+\delta\Psi^{-1})^{-1})\alpha^T\Psi^{-1}\alpha,$$

from which the result follows similarly to before. We obtain (5.8) using the lower bound in (5.7) and

$$\lambda_{\max}(\Psi(sm)^{-1}) = \lambda_{\max}((s\Psi(m))^{-1}) = \frac{1}{s}\lambda_{\max}(\Psi(m)^{-1}).$$ □

Existence of a minimizer. We summarize and derive important properties of the cost function J_δ, which follow mostly from the fact that J_δ is the variance of a BLUE. To avoid technical difficulties we assume that C is positive definite.

Lemma 5.9 (Properties of J_δ)**.** Let C be positive definite. The function J_δ has the following properties:

(i) $J_\delta : \mathbb{R}_{\geq 0}^{|K|} \to \mathbb{R}_{>0}$,

(ii) $J_\delta(m/\lambda) = \lambda J_{\delta/\lambda}(m)$ for all $\delta \geq 0$ and $\lambda > 0$,

(iii) Coupling of more models never increases the variance, that is for $k, j \in \{1, \dots, K\}$ with $S^k \subseteq S^j$ we have for all $\lambda \geq 0$ and $\delta > 0$

$$J_\delta(m + \lambda e_k) \geq J_\delta(m + \lambda e_j),$$

(iv) Increasing the number of samples never increases the variance, J_δ is monotonically decreasing in every argument,

(v) J_δ is twice continuously differentiable for $\delta > 0$ with derivatives

$$\begin{aligned} \partial_{m_k} J_\delta(m) &= -\alpha^T \Psi_\delta(m)^{-1} P^k (C^k)^{-1} R^k \Psi_\delta(m)^{-1} \alpha, \\ \partial_{m_k} \partial_{m_j} J_\delta(m) &= 2\alpha^T \Psi_\delta(m)^{-1} P^k (C^k)^{-1} R^k \Psi_\delta(m)^{-1} P^j (C^j)^{-1} R^j \Psi_\delta(m)^{-1} \alpha. \end{aligned}$$

(vi) J_δ is convex for $\delta > 0$,

(vii) J_δ is not strictly convex for $L \geq 3$.

Proof. The properties "(i)" and "(ii)" are straightforward to verify. For "(iii)" we use $S^k \subseteq S^j$, denote $I := S^j \setminus S^k$ with $|I| > 0$ and w.l.o.g. assume

$$C^j = \begin{pmatrix} C_{I,I} & C_{I,S^k} \\ C_{S^k,I} & C^k \end{pmatrix}.$$

Since the covariance matrix C^j is positive definite we use the inverse of the Schur complement [109] and the fact that the Schur complement is positive semi–definite

$$\begin{aligned} (C^j)^{-1} &= \begin{pmatrix} I & 0 \\ -(C^k)^{-1} C_{S^k,I} & I \end{pmatrix} \begin{pmatrix} (C^j/C^k)^{-1} & 0 \\ 0 & (C^k)^{-1} \end{pmatrix} \begin{pmatrix} I & -C_{I,S^k}(C^k)^{-1} \\ 0 & I \end{pmatrix} \\ &\geq \begin{pmatrix} I & 0 \\ -(C^k)^{-1} C_{S^k,I} & I \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & (C^k)^{-1} \end{pmatrix} \begin{pmatrix} I & -C_{I,S^k}(C^k)^{-1} \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & (C^k)^{-1} \end{pmatrix}. \end{aligned}$$

We conclude that for all $v \in \mathbb{R}^{|S^j|}$

$$v^T (C^j)^{-1} v \geq v_{S^k}^T (C^k)^{-1} v_{S^k}$$

and thus

$$P^k (C^k)^{-1} R^k \leq P^j (C^j)^{-1} R^j.$$

We use this to verify that for all $\lambda \geq 0$

$$\Psi(m + \lambda e_k) = \Psi(m) + \lambda P^k (C^k)^{-1} R^k \leq \Psi(m) + \lambda P^j (C^j)^{-1} R^j = \Psi(m + \lambda e_j).$$

For two symmetric positive definite matrices A, B it is well known that

$$A \leq B \quad \Leftrightarrow \quad A^{-1} \geq B^{-1},$$

which we use to conclude the statement

$$J_\delta(m+\lambda e_k) = \alpha^T(\Psi(m+\lambda e_k)+\delta I)^{-1}\alpha \geq \alpha^T((\Psi(m+\lambda e_j)+\delta I)^{-1}\alpha = J_\delta(m+\lambda e_j).$$

"(iv)": The computation is analogous to (iii), where we informally may use $S^k := \emptyset$ to directly obtain the result. Alternatively (v) shows that the first derivative is zero or negative.
"(v)": The expression for the first derivative follows from a standard derivative rule for inverse matrices [109]

$$\partial_{m_k}(\Psi_\delta(m))^{-1} = -\Psi_\delta(m)^{-1}(\partial_{m_k}\Psi_\delta(m))\Psi_\delta(m)^{-1}.$$

The expression for the second order derivative follows if we apply the derivative rule for the matrix product [109] for matrices A and B

$$\partial_{m_k}(A(m)B(m)) = (\partial_{m_k}A(m))B(m) + A(m)(\partial_{m_k}B(m)).$$

"(vi)": We verify that the Hessian is positive semi–definite. For $v \in \mathbb{R}^K$ we have

$$\begin{aligned} v^T H_{J_\delta} v &= \sum_{k,j=1}^{K} v_k 2\alpha^T \Psi_\delta(m)^{-1} P^k (C^k)^{-1} R^k \Psi_\delta(m)^{-1} P^j (C^j)^{-1} R^j \Psi_\delta(m)^{-1}\alpha v_j \\ &= 2\alpha^T \Psi_\delta(m)^{-1}\left(\sum_{k=1}^{K} v_k P^k (C^k)^{-1} R^k\right)\Psi_\delta(m)^{-1}\left(\sum_{j=1}^{K} v_j P^j (C^j)^{-1} R^j\right)\Psi_\delta(m)^{-1}\alpha \\ &= 2\alpha^T \Psi_\delta(m)^{-1}\Psi(v)\Psi_\delta(m)^{-1}\Psi(v)\Psi_\delta(m)^{-1}\alpha, \end{aligned}$$

which is always greater or equal to zero.
"(vii)": The matrix Ψ is symmetric as sum of $K = 2^L - 1$ symmetric matrices

$$\Psi(m) = \sum_{k=1}^{K} m_k P^k (C^k)^{-1} R^k.$$

The space of symmetric matrices has dimension $(L+1)L/2 < 2^L - 1 = K$ for $L \geq 3$. Therefore, there exists coefficients $\beta \neq 0$ such that

$$\Psi(\beta) = 0$$

and thus with $m_k = 1$ for all $k \in \{1, \ldots, K\}$ and for all λ sufficiently small

$$J_\delta(m) = \alpha^T(\Psi(m)+\delta I)^{-1}\alpha = \alpha^T(\Psi(m+\lambda\beta)+\delta I)^{-1}\alpha = J_\delta(m+\lambda\beta).$$

Therefore, J_δ is constant along the non–zero direction β and thus not strictly convex. □

We show that the relaxed sample allocation problem (5.6) has a minimizer, which is basically a result of J_δ being convex.

Theorem 5.10 (Existence of a minimizer)**.** Let C be positive definite and $\delta > 0$. Then there exists a minimizer of (5.6).

Proof. The function J_δ is convex due to Lemma 5.9. We optimize over the feasible region

$$\left\{ m_1, \dots, m_K \in \mathbb{R}^K \,|\, m_k \geq 0 \quad \text{for all } k \in \{1, \dots, K\}, \quad \sum_{k=1}^{K} m_k W^k = \mathbb{W}^{\text{budget}} \right\},$$

which is compact since $W^k > 0$ for all $k \in \{1, \dots, K\}$. All that remains to show is that this set is non–empty, which follows since it contains the MC estimator

$$m_k = \begin{cases} \frac{\mathbb{W}^{\text{budget}}}{W^k}, & \text{for } k \text{ with } S^k = \{1, \dots, L\}, \\ 0, & \text{otherwise.} \end{cases}$$

□

We write down the KKT conditions for the convex optimization problem (5.6). These are necessary and sufficient for the minimizers. Let $\xi^{\mathbb{W}}$ be the Lagrange multiplier associated with the cost constraint and $\xi_1, \dots, \xi_K$ with the positivity constraint. Then the minimizer satisfies

$$\begin{aligned} \alpha^T \Psi_\delta(m)^{-1} P^k (C^k)^{-1} R^k \Psi_\delta(m)^{-1} \alpha &= W^k \xi^{\mathbb{W}} - \xi_k && \text{for all } k \in \{1, \dots, K\}, \\ \sum_{k=1}^{K} m_k W^k &= \mathbb{W}^{\text{budget}}, \\ m_k &\geq 0 && \text{for all } k \in \{1, \dots, K\}, \\ m_k \xi_k &= 0 && \text{for all } k \in \{1, \dots, K\}. \end{aligned} \tag{5.9}$$

Sparsity of used model groups. The minimization problem (5.6) has $K = 2^L - 1$ variables, which is exponential in L. Furthermore, a practical implementation requires us to ceil the number of samples m to be integer. The additional cost for this operation is bounded by an additional evaluation of every model group

$$\sum_{k=1}^{2^L - 1} W^k, \tag{5.10}$$

which is a bound that is also exponential in L. Fortunately, we are always able to reduce the number of used model groups to at most L without increasing the variance or the cost. We formally denote this as $|U_S(m)| \leq L$, where we make the dependence of the used model groups $U_S(m)$ on the number of samples m explicit.

Theorem 5.11 (Sparse solution)**.** Let C be positive definite, $\delta > 0$ and m be a feasible sample allocation of (5.6) with $|U_S(m)| > L$. Then there exists a feasible point m' with $|U_S(m')| \leq L$ and

$$J_\delta(m') \leq J_\delta(m).$$

In particular, there exists a minimizer m^* of (5.6) with $|U_S(m^*)| \leq L$.

Proof. Let m be a feasible point such that w.l.o.g. $m_1, \dots, m_{L+1} > 0$. We now construct a direction along which the variance J_δ is constant and the cost does not increase. For $x := \Psi_\delta(m)^{-1} \alpha$ there exists $\beta \in \mathbb{R}^{L+1} \setminus \{0\}$ such that the following linear combination is zero

$$\sum_{\ell=1}^{L+1} \beta_\ell R^\ell (C^\ell)^{-1} P^\ell x = \sum_{\ell=1}^{L+1} \beta_\ell x^\ell = 0, \tag{5.11}$$

which follows from a dimension counting argument of the vectors $x^1, \dots, x^{L+1} \in \mathbb{R}^L$ with $x^\ell := R^\ell (C^\ell)^{-1} P^\ell x$. Furthermore, we demand that these coefficients satisfy

$$\sum_{\ell=1}^{L+1} \beta_\ell W^\ell \leq 0, \tag{5.12}$$

which can always be achieved by suitably changing the sign of β, that is redefining β as $-\beta$ if necessary. For notational purposes we now view β as a vector in $\mathbb{R}^K$ where we set $\beta_k := 0$ for $k \in \{L+2, \dots, K\}$. We define the maximum scale parameter $s_{\max} \in \mathbb{R}$ such that we are at the boundary of the feasible region with the ray originating in m in direction of β

$$s_{\max} := \max\{s \geq 0 \,|\, m_\ell + s\beta_\ell \geq 0 \text{ for all } \ell \in \{1, \dots, L+1\}\}. \tag{5.13}$$

The combination of $\beta \neq 0$ and $W^\ell > 0$ with (5.12) shows that there exists a negative coefficient $\beta_\ell < 0$ and thus $s_{\max} < +\infty$ is well defined. We verify that the sample allocation $m' := m + s_{\max}\beta \geq 0$ uses at least one less model group, has smaller or equal cost and equal variance compared to m.

- $m' \geq 0$: By definition of β we have $m'_k = m_k \geq 0$ for $k \in \{L+2, \dots, K\}$. For $\ell \in \{1, \dots, L+1\}$ we have $m'_\ell = m_\ell + s_{\max}\beta_\ell \geq 0$ from the definition of $s_{\max}$ in (5.13).

- m' uses at least on less model group: By definition of $s_{\max}$ there exists an index $\ell \in \{1, \dots, L+1\}$ such that $m_\ell + s_{\max}\beta_\ell = 0$. Therefore, m' does not use S^ℓ and no other additional model groups are used for m', hence $|U_S(m')| < |U_S(m)|$.

- m' has smaller or equal cost: This is a consequence of (5.12) and $s_{\max} \geq 0$

$$\sum_{k=1}^{K} (m_k + s_{\max}\beta_k) W^k = \sum_{k=1}^{K} m_k W^k + s_{\max} \sum_{\ell=1}^{L+1} \beta_\ell W^\ell \leq \sum_{k=1}^{K} m_k W^k.$$

- $J_\delta(m') = J_\delta(m)$: The definition of $x := \Psi_\delta(m)^{-1}\alpha$ shows

$$\alpha = \Psi_\delta(m)x = \sum_{\ell=1}^{L+1} m_\ell R^\ell (C^\ell)^{-1} P^\ell x + \sum_{k=L+2}^{K} m_k R^k (C^k)^{-1} P^k x + \delta x.$$

  We now use (5.11) to insert zero into the first term

$$\begin{aligned} \alpha &= \sum_{\ell=1}^{L+1} (m_\ell + s_{\max}\beta_\ell) R^\ell (C^\ell)^{-1} P^\ell x + \sum_{k=L+2}^{K} m_k R^k (C^k)^{-1} P^k x + \delta x \\ &= (\Psi(m + s_{\max}\beta) + \delta I)x \\ &= \Psi_\delta(m')x. \end{aligned}$$

  The matrix $\Psi_\delta(m')$ is positive definite and thus the variances are equal

$$J_\delta(m) = \alpha^T \Psi_\delta(m)^{-1}\alpha = \alpha^T x = \alpha^T \Psi_\delta(m')^{-1}\alpha = J_\delta(m').$$

We now inductively repeat the process outlined in this proof for m' until we reach a sample allocation $\overline{m}$ such that the initial assumption $|U_S(\overline{m})| > L$ is false. Hence there exists a sample allocation $\overline{m}$ with

$$\begin{aligned} |U_S(\overline{m})| &\leq L, \\ J_\delta(\overline{m}) &= J_\delta(m), \\ \sum_{k=1}^{K} \overline{m}_k W^k &\leq \sum_{k=1}^{K} m_k W^k = \mathbb{W}^{\text{budget}}, \\ m_k &\geq 0 \qquad \text{for all } k \in \{1, \dots, K\}. \end{aligned}$$

The sample allocation $\overline{m}$ is infeasible if it is truly cheaper than m. We fix this by scaling the number of samples such that for another sample allocation $\widetilde{m}$

$$\widetilde{m}_k := \overline{m}_k \frac{\mathbb{W}^{\text{budget}}}{\sum_{k=1}^{K} \overline{m}_k W^k} \geq \overline{m}_k \qquad \text{for all } k \in \{1, \dots, K\}.$$

The cost constraint is now satisfied with equality and since the variance J_δ is monotonically decreasing in every argument due to Lemma 5.9 (iv)

$$J_\delta(\widetilde{m}) \leq J_\delta(\overline{m}) = J_\delta(m). \qquad \square$$

It is now trivial to find an upper bound on the additional cost introduced from the ceiling since every model is evaluated additionally at most L times, which is a substantial improvement over (5.10)

$$L \sum_{\ell=1}^{L} w_\ell.$$

However, in practice this bound is often better since expensive models often occur only in few of the used model groups. We remark that ceiling the number of samples does not necessarily lead to the SAOB. However, if the budget $\mathbb{W}^{\text{budget}}$ is large or δ is small, then it is straightforward to argue that the error or the additional cost can be neglected.
We are now in the position to provide an example where the optimal sample allocation in (5.6) is not unique.

Example 5.12 (Non–uniqueness of an optimal solution)**.** We know that the MFMC estimator is the BLUE for $L \leq 2$ and that the MC estimator is the BLUE for $L = 1$. We show that if the MC estimator has the same variance as the MFMC estimator, then the optimal sample allocation for the BLUE is not unique. For $\alpha = (0, 1)^T$ we narrow down the possible BLUE by examining which models groups we are able to use together.

- $\{1\}$: Impossible, bias constraint $\beta^1_{[1]} = 0$, additional costs to evaluate Z_1,
- $\{2\}$: Possible, the MC estimator,
- $\{1, 2\}$: Impossible, bias constraint $\beta^3_{[1]} = 0$, additional costs to evaluate Z_1,
- $\{1\}, \{2\}$: Impossible, bias constraints $\beta^1_{[1]} = 0$, additional costs to evaluate Z_1,
- $\{1\}, \{1, 2\}$: Possible, the MFMC estimator,
- $\{1, 2\}, \{2\}$: Impossible, bias constraints $\beta^3_{[1]} = 0$, additional costs to evaluate Z_1,

- $\{1\}, \{2\}, \{1,2\}$: Possible, combination of the MC and MFMC estimator.

If the last combination $\{1\}, \{2\}, \{1,2\}$ has the best sample allocation, then we are able to use Theorem 5.11 to reduce the number of active model groups to two or less, which is the MFMC or the MC estimator. Hence, it suffices to construct an example where both have the same variance. We use Theorem 3.26 and require the data to satisfy

$$\mathbb{V}\big[\widehat{\mu}_2^{\mathrm{MFMC}}\big] = \frac{\sigma_2^2}{\mathbb{W}^{\mathrm{budget}}}\Big(\big(w_1\big(\rho_{2,1}^2 - 0\big)\big)^{1/2} + \big(w_2\big(\rho_{2,2}^2 - \rho_{2,1}^2\big)\big)^{1/2}\Big)^2 = \frac{w_2}{\mathbb{W}^{\mathrm{budget}}}\sigma_2^2 = \mathbb{V}\big[\widehat{\mu}_2^{\mathrm{MC}}\big].$$

We choose $\rho_{2,1}^2 = 1/2$ and $w_2 = 1$ such that w_1 has to satisfy

$$\left(w_1^{1/2}\left(\frac{1}{2}\right)^{1/2} + \left(\frac{1}{2}\right)^{1/2}\right)^2 = 1,$$

which is true for $w_1 = (\sqrt{2}-1)^2$. With these choices for w_1, w_2 and $\rho_{2,1}$ the assumptions of Theorem 3.26 are satisfied and thus our derivation is valid. ◇

5.3 First sample allocation, then BLUE

Sample allocation. In the previous section we derived some properties of the optimal sample allocation by adding the matrix δI which we interpreted as additional model evaluations. This change was made purely because of technical reasons to avoid the difficulty of the matrix inversion. We now derive a stronger result for $\delta = 0$ by first optimizing over m and then over the coefficients β for the variance of a linear unbiased estimator

$$\mathbb{V}[\widehat{\mu}_\alpha] = \sum_{k=1}^{K} \frac{(\beta^k)^T C^k \beta^k}{m_k}.$$

In the first step β is not necessarily such that $\widehat{\mu}_\alpha$ is a BLUE. We have to make sure that the cost constraint is satisfied and we do not divide by zero. The case $\beta^k = 0$ is equivalent to not using the model group S^k and thus we w.l.o.g. may define $m_k = 0$. On the other hand if $\beta^k \neq 0$ we w.l.o.g. assume that $m_k > 0$. The set of used model groups is then entirely defined by the coefficients β

$$U_\beta := \{k \in \{1, \dots, K\} \,|\, (\beta^k)^T C^k \beta^k > 0\}.$$

For the special case of a positive definite covariance matrix C, we have

$$U_\beta = \{k \in \{1, \dots, K\} \,|\, \beta^k \neq 0\} = U_S.$$

We write down the relaxed sample allocation problem for a linear unbiased estimator with fixed coefficient β

$$\begin{aligned} \min_{m_1,\dots,m_K \in \mathbb{R}} \quad & J(m) := \sum_{k \in U_\beta} \frac{(\beta^k)^T C^k \beta^k}{m_k} \\ \text{such that} \quad & \sum_{k=1}^{K} m_k W^k = \mathbb{W}^{\mathrm{budget}}, \\ & m_k \geq 0 \qquad \text{for all } k \in \{1, \dots, K\}. \end{aligned} \tag{5.14}$$

We remark that the MLMC estimator is a linear unbiased estimator and thus with

$$\begin{aligned} S^1 &:= \{1\}, & \beta^1 &:= 1, \\ S^2 &:= \{1,2\}, & \beta^2 &:= (-1,1), \\ &\vdots & &\vdots \\ S^L &:= \{L-1, L\}, & \beta^L &:= (-1,1), \end{aligned}$$

the cost function J is

$$J(m) = \sum_{\ell=1}^{L} \frac{\mathbb{V}[Z_\ell - Z_{\ell-1}]}{m_\ell}.$$

This particular instance of (5.14) is exactly the sample allocation problem for the MLMC estimator, which has a unique solution according to Theorem 3.46. We show that this is in general the case for arbitrary β in (5.14) and unsurprisingly the exact statement as well as its proof are quite similar.

Theorem 5.13 (Optimal sample allocation for linear unbiased estimators)**.**
Let $(\beta^k)^T C^k \beta^k = 0$ for all $k \in \{1, \dots, K\}$. Then any feasible point of (5.14) is optimal with $J(m) = 0$. If $(\beta^k)^T C^k \beta^k > 0$ for some $k \in \{1, \dots, K\}$, then the unique optimal sample allocation of (5.14) is

$$m_k = \frac{\mathbb{W}^{\text{budget}}}{\sum_{k \in U_\beta} [(\beta^k)^T C^k \beta^k W^k]^{1/2}} \left(\frac{(\beta^k)^T C^k \beta^k}{W^k} \right)^{1/2} \qquad \text{for all } k \in \{1, \dots, K\}. \tag{5.15}$$

The variance at this minimizer is

$$J(m) = \frac{1}{\mathbb{W}^{\text{budget}}} \left(\sum_{k \in U_\beta} \left[(\beta^k)^T C^k \beta^k W^k \right]^{1/2} \right)^2. \tag{5.16}$$

Proof. First, observe that the statement of this theorem for$(\beta^k)^T C^k \beta^k = 0$ for all $k \in \{1, \dots, K\}$ is trivial. Thus assume now that $(\beta^k)^T C^k \beta^k > 0$ for some k, which immediately shows that the denominator in (5.15) is well defined. Notice that for k with $(\beta^k)^T C^k \beta^k = 0$ the choice $m_k = 0$ is optimal since otherwise this incurs a cost proportional to $W^k > 0$ but does not decrease the variance. Thus (5.15) is valid for those model groups. Since we divide by m_k in J there exists a positive constant $c > 0$ such that the number of samples is bounded from below

$$m_k \geq c \qquad \text{for all } k \in U_\beta.$$

Similarly, the cost constraint and $W^k > 0$ ensures that there exists a second positive constant $c > 0$ such that the number of samples is bounded from above

$$m_k \leq c \qquad \text{for all } k \in U_\beta.$$

J is convex and we optimize over a compact and non–empty set, hence there exists a solution of (5.14). We write down the necessary and sufficient KKT conditions with Lagrange-multiplier $\xi^{\mathbb{W}} \in \mathbb{R}$ associated to the cost and $\xi_k \in \mathbb{R}$ for $k \in U_\beta$ associated with

the positivity constraint

$$-\frac{(\beta^k)^T C^k \beta^k}{m_k^2} + \xi^{\mathbb{W}} W^k - \xi_k = 0 \qquad \text{for all } k \in U_\beta, \tag{5.17}$$

$$\sum_{k \in U_\beta} m_k W^k = \mathbb{W}^{\text{budget}}, \tag{5.18}$$

$$m_k \geq 0 \qquad \text{for all } k \in U_\beta, \tag{5.19}$$

$$\xi_k \geq 0 \qquad \text{for all } k \in U_\beta, \tag{5.20}$$

$$m_k \xi_k = 0 \qquad \text{for all } k \in U_\beta. \tag{5.21}$$

We have already verified that m_k is positive for all $k \in U_\beta$ and thus $\xi_k = 0$. Furthermore, since there exists a solution we conclude $\xi^{\mathbb{W}} > 0$, otherwise (5.17) cannot be satisfied. We reformulate this equation to arrive at

$$m_k = \left(\frac{(\beta^k)^T C^k \beta^k}{\xi^{\mathbb{W}} W^k} \right)^{1/2}$$

and insert this into (5.18)

$$(\xi^{\mathbb{W}})^{1/2} = \frac{\sum_{k \in U_\beta} \left[(\beta^k)^T C^k \beta^k W^k \right]^{1/2}}{\mathbb{W}^{\text{budget}}}.$$

We combine the last two equations to obtain (5.15) and a straightforward computation shows (5.16). □

We now choose the coefficients β to minimize the variance and respecting the bias constraint. We achieve this if we minimize over the inner expression of (5.16)

$$\begin{aligned} \min_{\beta} \qquad & J(\beta) := \sum_{k=1}^{K} \left[(\beta^k)^T C^k \beta^k W^k \right]^{1/2} \\ \text{such that} \qquad & \sum_{k=1}^{K} P^k \beta^k = \alpha. \end{aligned} \tag{5.22}$$

The cost function J from (5.16) now includes values that are zero $\beta^k = 0$. We remark that we now do not have any issue stating this problem w.r.t. the matrix inversion of Ψ as this was the case in Section 5.2 where we have added the artificial matrix δI.

Theorem 5.14 (Existence of optimal coefficients)**.** There exists a solution of (5.22).

Proof. First, we assume that C is positive definite. Then since $W^k > 0$ we define the norms $\| \cdot \|_k$ for all $k \in \{1, \dots, K\}$

$$\|v\|_k := \left[(v)^T W^k C^k v \right]^{1/2} \qquad \text{for all } v \in \mathbb{R}^{|S^k|}.$$

The cost function is a sum over different norms and thus convex

$$J(\beta) = \sum_{k=1}^{K} \|\beta^k\|_k.$$

The minimization problem (5.22) has a feasible point, since the bias constraint is satisfied for $\beta^L = \alpha$ for $S^L := \{1, \dots, L\}$ and $\beta^k = 0$ for $k \neq L$. We further conclude that for every $k \in \{1, \dots, K\}$

$$\lim_{\|\beta^k\| \to +\infty} J(\beta) = +\infty.$$

We are thus able to w.l.o.g. restrict the set of possible coefficients to a bounded set

$$\left\{ \beta \,|\, \sum_{k=1}^{K} P^k \beta^k = \alpha, \quad \|\beta^k\| \leq c \quad \text{for all } k \in \{1, \dots, K\} \right\},$$

where $c > 0$ is a sufficiently large positive constant. A minimizer now exists, since J is continuous and we optimize over a non–empty compact set.
Now let C be non–negative definite. We split $\mathbb{R}^{|S^k|} = X^k + Y^k$ such that C^k is positive definite on X^k and Y^k is the space of eigenvectors with zero eigenvalue. We rewrite (5.22) such that

$$\begin{aligned} \min_{\substack{x^1 \in X^1, \dots, x^K \in X^K, \\ y^1 \in Y^1, \dots, y^K \in Y^K}} \quad & J(\beta) = \sum_{k=1}^{K} \|x^k\|_k \\ \text{such that} \quad & \sum_{k=1}^{K} P^k x^k = \alpha - \sum_{k=1}^{K} P^k y^k. \end{aligned}$$

The vectors y^k do not influence the value of J which allows us to replace the bias constraint with a constraint in a smaller space

$$P^{Y^\perp} \sum_{k=1}^{K} P^k x^k = P^{Y^\perp} \alpha,$$

where $P^{Y^\perp}$ is the projection onto the space $Y^\perp$

$$Y^\perp := \{ v \in \mathbb{R}^{|S^k|} \,|\, (v, P^k y^k) = 0 \quad \text{for all } y^k \in Y^k, \quad k \in \{1, \dots, K\} \}.$$

The resulting sample allocation problem now reads

$$\begin{aligned} \min_{x^1 \in X^1, \dots, x^K \in X^K} \quad & J(\beta) = \sum_{k=1}^{K} \|x^k\|_k \\ \text{such that} \quad & P^{Y^\perp} \sum_{k=1}^{K} P^k x^k = P^{Y^\perp} \alpha. \end{aligned}$$

The function $\|\cdot\|_k$ is a norm on X^k, hence similarly to the case of positive definite C the existence of a minimizer follows. □

We remark that an optimizer in Theorem 5.14 corresponds to an optimizer in Theorem 5.10 if $\delta = 0$ and vice versa. Furthermore, it is not straightforward to actually compute an optimal coefficient β, since the norm $\|\cdot\|_k$ is not differentiable at zero which requires an expensive case distinction w.r.t. the used model groups.
Let us now assume that we have computed the optimal coefficients β. We are able to compute the number of samples from (5.14)

$$m_k = \frac{\mathbb{W}^{\text{budget}}}{\sum_{k=1}^{K} [(\beta^k)^T C^k \beta^k W^k]^{1/2}} \left(\frac{(\beta^k)^T C^k \beta^k}{W^k} \right)^{1/2}, \tag{5.23}$$

however this expression is only valid if C is positive definite. It is straightforward to verify that if the denominator is zero, then we found an estimator with zero variance and thus the choice $m = 0$ seems reasonable. However, this violates the existence of a linear unbiased estimator with bias $\alpha \neq 0$. This fact even holds if the denominator is not zero, which we show in a continuation of Example 5.5.

Example 5.15 (C not invertible)**.** Let us assume that $L := 2$, $w_1 := w_2 := 1$, $\mathbb{W}^{\text{budget}} := 1$, $\alpha := (1,1)^T$ and the covariance matrix $C := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. With $S^1 := \{1\}, S^2 := \{2\}$ and $S^3 := \{1,2\}$ the cost function and bias constraint is

$$J(\beta) = \|\beta^1\|_1 + \|\beta^2\|_2 + \|\beta^3\|_3 = |\beta_1^2| + 2|\beta_2^3|,$$
$$\alpha = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \beta_1^1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \beta_1^2 \end{pmatrix} + \begin{pmatrix} \beta_2^3 \\ \beta_2^3 \end{pmatrix}.$$

The minimizer is clearly $\beta_1^1 = \beta_1^2 = 1$ and zero for all other values. We compute the number of samples with (5.23)

$$m_1 = 0, \quad m_2 = 1, \quad m_3 = 0.$$

However, since $U_\alpha = \{1,2\} \neq U_Z = \{2\}$ there exists no linear unbiased estimator with this sample allocation. ◇

We are able to circumvent the problems of the previous example if we require a single evaluation of the model group U_α. This step can be postponed after computing the optimal coefficients β. We further have to ceil the number of samples in (5.23) to obtain a reasonable approximation of the SAOB. In this sense, we removed both the need for adding the additional matrix δI as well as the assumption that C is positive definite.

Sparsity of used model groups. We provide a proof of the existence of a sparse solution β that satisfies $|U_\beta| \leq L$. This is a consequence of the sum of norms property, since we may view J as ℓ_1-norm over the different coefficients β^k in the following sense

$$J(\beta) = \|\beta\|_{\ell_1} := \sum_{k=1}^{K} \|\beta^k\|_k.$$

We keep the notation of $\|\cdot\|_k$ even though this function may only be a semi–norm if C is not positive definite. We never implicitly rely on its norm properties. The proof of the sparsity result is in spirit similar to the proof of Theorem 5.11.

Theorem 5.16 (Sparse solution)**.** Let β be a feasible point of (5.22) with $|U_\beta| > L$. Then there exists a feasible point β' of (5.22) with $|U_{\beta'}| \leq L$ and $J(\beta') \leq J(\beta)$. In particular, there exists a minimizer β^* of (5.22) with $|U_{\beta^*}| \leq L$.

Proof. We construct a direction along which the bias remains unchanged and the variance does not increase. For a feasible point β with $|U_\beta| > L$ we w.l.o.g. assume $\beta^1, \dots, \beta^{L+1} \neq 0$. Then by a dimension counting argument applied to $P^\ell \beta^\ell \in \mathbb{R}^L$ there exists a linear combination $v \in \mathbb{R}^{L+1} \setminus \{0\}$ such that

$$\sum_{\ell=1}^{L+1} v_\ell P^\ell \beta^\ell = 0.$$

For notational purposes we now view v as element of $v \in \mathbb{R}^K$, where we extend this vector by zeros $v_k = 0$ for $k \in \{L+2, \dots, K\}$. The bias does not change along this direction since for arbitrary $s \in \mathbb{R}$

$$\alpha = \sum_{k=1}^{K} P^k \beta^k = \sum_{k=1}^{L+1} (1 + sv_k) P^k \beta^k + \sum_{k=L+2}^{K} (1 + s0) P^k \beta^k = \sum_{k=1}^{K} P^k [(1 + sv_k)\beta^k].$$

We write down the value of the cost function

$$\begin{aligned} J((1+sv_1)\beta^1, \dots, (1+sv_K)\beta^K) &= \sum_{k=1}^{K} \|(1+sv_k)\beta^k\|_k = \sum_{k=1}^{K} |(1+sv_k)| \|\beta^k\|_k \\ &= \sum_{\ell=1}^{L+1} |(1+sv_\ell)| \|\beta^\ell\|_\ell + \sum_{k=L+2}^{K} \|\beta^k\|_k. \end{aligned}$$

Only the first sum depends on s and we now show that there exists an index $\ell \in \{1, \dots, L+1\}$ such that we are able to choose $s = -1/v_\ell$, which removes the ℓ–th model group. We view J as function of s and compute the derivative for s close to zero

$$\partial_s J(s) := \partial_s J((1+sv_1)\beta^1, \dots, (1+sv_K)\beta^K) = \sum_{\ell=1}^{L+1} v_\ell \|\beta^\ell\|_\ell. \tag{5.24}$$

This expression is valid only for small s such that $1 + sv_\ell > 0$ for all ℓ since the modulus $|\cdot|$ is not differentiable at zero. The key idea is to recognize that J as function of s is locally linear near 0 and globally continuous, which allows to compute its exact value using only the first derivative. We then choose s such that $1 + sv_\ell \geq 0$ with equality for at least one model group. We distinguish between three cases.

- $\partial_s J < 0$: We choose s as maximum

  $$s_{\max} := \max\{s \geq 0 \,|\, 1 + sv_\ell \geq 0, \quad \text{for all } \ell \in \{1, \dots, L+1\}\}.$$

  From (5.24) together with $\partial_s J < 0$ we conclude that there exists $\ell \in \{1, \dots, L+1\}$ with $v_\ell < 0$ and thus $s_{\max} < +\infty$ is well defined. Since $\partial_s J < 0$, J is locally linear and $s_{\max} > 0$ we conclude that

  $$J((1 + s_{\max} v_1)\beta^1, \dots, (1 + s_{\max} v_K)\beta^K) < J(\beta^1, \dots, \beta^K).$$

  Moreover, by definition of $s_{\max}$ and with β' such that $\beta'^{,k} := (1 + s_{\max} v_k)\beta^k$ for all $k \in \{1, \dots, K\}$ we conclude $|U_{\beta'}| < |U_\beta|$, since at least one less model group $\ell \in \{1, \dots, L+1\}$ is used. In this case, we have $v_\ell \neq 0$, $1 + s_{\max} v_\ell = 0$ and thus $s_{\max} = -1/v_\ell$.

- $\partial_s J > 0$: The proof is analogous to $\partial_s J > 0$. Here we have to choose $s_{\min}$ as minimal s such that $1 + sv_\ell \geq 0$ for all $\ell \in \{1, \dots, L+1\}$. Then we have $0 > s_{\min} > -\infty$ and the result follows.

- $\partial_s J = 0$: In this case the variance is constant and we may choose $s_{\max,\min}$ either as maximum or minimum s such that $1 + sv_\ell \geq 0$ for all $\ell \in \{1, \dots, L+1\}$ with equality for at least one model group. Once again from $v \neq 0$ we are able to conclude $-\infty < s_{\max,\min} < +\infty$ and thus

  $$J((1 + s_{\max,\min} v_1)\beta^1, \dots, (1 + s_{\max,\min} v_K)\beta^K) = J(\beta^1, \dots, \beta^K).$$

  The definition of β' is similar to before.

We started from coefficients β with $|U_\beta| > L$ and constructed other coefficients β' such that

$$\begin{aligned} J(\beta') &\leq J(\beta), \\ \sum_{k=1}^{K} P^k \beta'^{,k} &= \sum_{k=1}^{K} P^k \beta^k = \alpha, \\ |U_{\beta'}| &< |U_\beta|. \end{aligned}$$

We now inductively continue the procedure outlined in this proof until we obtain the coefficients that satisfy the statement of the theorem. □

We now return to Example 5.12 and verify that we are actually in the case $\partial_s J = 0$.

Example 5.17 (Non–uniqueness of SAOB)**.** The MC estimator and the MFMC estimator have the same variance if $\rho_{2,1} = \sqrt{1/2}$, $w_1 = (\sqrt{2}-1)^2$ and $w_2 = 1$. We assume $\sigma_1^2 = \sigma_2^2 = 1$. The coefficients for the MC estimator are $\beta^2 = 1$ and zero otherwise. We compute the respective value of the norm

$$\|\beta^2\|_2 = \sqrt{(\beta^2)^T C^2 \beta^2 W^2} = \sqrt{w_2} = 1.$$

We now compute the respective coefficients of the MFMC estimator. This estimator reads

$$\begin{aligned} \widehat{\mu}^{\mathrm{MFMC}} &= \frac{1}{m_3}\sum_{i=1}^{m_3} Z_2^i - \rho_{2,1}\left(\frac{1}{m_3}\sum_{i=1}^{m_3} Z_1^i - \frac{1}{m_1+m_3}\sum_{i=1}^{m_1+m_3} Z_1^i\right) \\ &= \frac{1}{m_3}\sum_{i=1}^{m_3} Z_2^i - \rho_{2,1}\frac{m_1}{m_1+m_3}\left(\frac{m_1+m_3}{m_1}\left[\frac{1}{m_3} - \frac{1}{m_1+m_3}\right]\sum_{i=1}^{m_3} Z_1^i - \frac{1}{m_1}\sum_{i=m_3+1}^{m_1+m_3} Z_1^i\right) \\ &= \frac{1}{m_3}\sum_{i=1}^{m_3} Z_2^i - \rho_{2,1}\frac{m_1}{m_1+m_3}\left(\frac{1}{m_3}\sum_{i=1}^{m_3} Z_1^i - \frac{1}{m_1}\sum_{i=m_3+1}^{m_1+m_3} Z_1^i\right). \end{aligned}$$

We insert the optimal number of samples (3.27) (where the m_ℓ^* denotes the total number of samples) to conclude

$$\begin{aligned} \beta_1^3 = -\beta^1 = -\rho_{2,1}\frac{m_1}{m_1+m_3} &= -\rho_{2,1}\frac{\left(\frac{\rho_{2,1}^2-\rho_{2,0}^2}{w_1}\right)^{1/2} - \left(\frac{\rho_{2,2}^2-\rho_{2,1}^2}{w_2}\right)^{1/2}}{\left(\frac{\rho_{2,1}^2-\rho_{2,0}^2}{w_1}\right)^{1/2}} = -\rho_{2,1}(1-w_1^{1/2}) \\ &= -(\sqrt{2}-1) = -\sqrt{w_1}. \end{aligned}$$

For the MFMC estimator we thus have $\beta^3 = (-\sqrt{w_1}, 1)^T$ and $\beta^1 = \sqrt{w_1}$. We compute the respective norms

$$\begin{aligned} \|\beta^1\|_1 &= \sqrt{(\beta^1)^T C^1 \beta^1 W^1} = \sqrt{w_1 w_1} = w_1, \\ \|\beta^3\|_3 &= \sqrt{(\beta^3)^T C^3 \beta^3 W^3} = \sqrt{(w_1 - 2\sqrt{w_1}\rho_{2,1} + 1)(w_1+1)} = 2\sqrt{2} - 2 = 1 - w_1. \end{aligned}$$

We furthermore compute the linear combination v such that

$$0 = \sum_{\ell=1}^{3} v_\ell P^\ell \beta^\ell = v_1\begin{pmatrix}\sqrt{w_1}\\0\end{pmatrix} + v_2\begin{pmatrix}0\\1\end{pmatrix} + v_3\begin{pmatrix}-\sqrt{w_1}\\1\end{pmatrix},$$

which is satisfied for $v_1 = v_3 = 1$ and $v_2 = -1$. Now observe that the expression for the derivative of J w.r.t. the scaling s in (5.24) is

$$\partial_s J(s) = \sum_{\ell=1}^{3} v_\ell \|\beta^\ell\|_\ell = (w_1) - (1) + (1 - w_1) = 0.$$

We thus conclude that for all $\lambda \in [0, 1]$ the following values of β lead to optimal J under the bias constraint

$$\begin{aligned} \beta^1 &= 0 + \lambda\sqrt{w_1}, \\ \beta^2 &= 1 - \lambda, \\ \beta^3 &= 0 + \lambda\begin{pmatrix} -\sqrt{w_1} \\ 1 \end{pmatrix}, \end{aligned} \tag{5.25}$$

where $\lambda = 0$ is the MC estimator and $\lambda = 1$ the MFMC estimator. For $\lambda \in (0, 1)$ we have a convex combination of these two estimators. ◇

Lower variance bound. Let us derive the lower variance from the perspective of (5.15) and (5.16). We assume that the high fidelity model is in the model group S^1 and we decrease the cost of $Z_1, \dots, Z_{L-1}$ to zero $w_1, \dots, w_{L-1} \to 0$ to conclude

$$\begin{aligned} m_k &\to +\infty \qquad \text{for all } k \in \{2, \dots, K\}, \\ m_1 &\to 1, \\ J(\beta) &\to \frac{1}{\mathbb{W}^{\text{budget}}} (\beta^1)^T C^1 \beta^1 w_L = (\beta^1)^T C^1 \beta^1, \end{aligned}$$

where we have used $\mathbb{W}^{\text{budget}} = w_L$. The lower variance bound for the SAOB and thus for any linear unbiased estimator is

$$\mathbb{V}^{\min}[\widehat{\mu}_L^{\text{SAOB}}] = \mathbb{V}^{\min}_{S^1 \setminus \{L\}} = \mathbb{V}^{\min} = \min_\beta \quad (\beta^1)^T C^1 \beta^1,$$

where we have to satisfy the bias constraint in the last minimization problem

$$e_L = \alpha = \sum_{k=1}^{K} P^k \beta^k.$$

This is exactly the bound from Corollary 4.18.

5.4 Characterisation of the set of minimizers

Example 5.17 shows that the set of optimizers can be described as the convex hull of finitely many estimators with at most L active model groups. We verify that this is always the case if C is positive definite. Let us define the set

$$\{P^k \beta^k \mid k \in U_\beta\}. \tag{5.26}$$

This section proceeds with the following three steps:

1. Every minimizer is a convex combination of minimizers where the vectors (5.26) are linear independent.

2. If the vectors (5.26) are linear independent for a minimizer β, then there is no other minimizer β' such that $U_\beta = U_{\beta'}$. In this sense β is unique.

3. Therefore, every minimizer is a convex combination of finitely many minimizers where the vectors (5.26) are linearly independent and thus necessarily uses at most L model groups.

We start with the first claim.

Lemma 5.18 (Convex combination of minimizers)**.** The minimizer β of (5.22) is a convex combination of minimizers $\beta^{[1]}, \dots, \beta^{[N]}$ of (5.22)

$$\beta = \sum_{j=1}^{N} \lambda_j \beta^{[j]} \qquad \text{for some } \lambda_1, \dots, \lambda_N \in [0,1] \text{ with } \sum_{j=1}^{N} \lambda_j = 1,$$

where the vectors (5.26) are linear independent for each $\beta^{[j]}$.

Proof. Let β be a minimizer for which (5.26) is not linear independent such that

$$V_\beta := \left\{ v \in \mathbb{R}^K \mid \sum_{k \in U_\beta} v_k P^k \beta^k = 0, \quad v_k = 0 \text{ for } k \notin U_\beta \right\}$$

has dimension $d_\beta := \dim(V_\beta) > 0$. We show that β is a convex combination of two other minimizers $\overline{\beta}$ and $\widetilde{\beta}$ with $d_{\widetilde{\beta}} < d_\beta$ and $d_{\overline{\beta}} < d_\beta$. We then repeat this argument for both $\widetilde{\beta}$ and $\overline{\beta}$ until we obtain the minimizers $\beta^{[1]}, \dots, \beta^{[N]}$ for which $d_{\beta^{[j]}} = 0$ and thus the vectors (5.26) are linearly independent. Since a convex combination of a convex combination is again a convex combination, the minimizer β must be a convex combination of $\beta^{[1]}, \dots, \beta^{[N]}$ and thus the statement of the theorem holds. Therefore, we only have to verify that β is a convex combination of minimizers $\widetilde{\beta}$ and $\overline{\beta}$ with $d_{\widetilde{\beta}} < d_\beta$ and $d_{\overline{\beta}} < d_\beta$. In the proof of Theorem 5.16, whose notation we follow, there exists a vector v such that for all s moving along the direction

$$(1 + sv_1)\beta^1, \dots, (1 + sv_K)\beta^K$$

does not change the bias. Since β is a minimizer, we are in the setting $\partial_s J(s) = 0$ of Theorem 5.16, since otherwise we are able to find a feasible point with smaller value of J. Then there exists both $s_{\max} > 0$ and $s_{\min} < 0$ and two distinct minimizers

$$\widetilde{\beta} := ((1 + s_{\max} v_1)\beta^1, \dots, (1 + s_{\max} v_K)\beta^K)^T,$$
$$\overline{\beta} := ((1 + s_{\min} v_1)\beta^1, \dots, (1 + s_{\min} v_K)\beta^K)^T.$$

Each of these uses at least one less model group compared to β, hence $d_{\widetilde{\beta}} < d_\beta$ and $d_{\overline{\beta}} < d_\beta$. The minimizer β is a convex combination of $\widetilde{\beta}$ and $\overline{\beta}$, since for $\lambda = -s_{\min}/(s_{\max} - s_{\min}) \in [0,1]$

$$\lambda \widetilde{\beta}^k + (1-\lambda)\overline{\beta}^k = -\frac{s_{\min}}{s_{\max} - s_{\min}}(1 + s_{\max} v_k)\beta^k + \frac{s_{\max}}{s_{\max} - s_{\min}}(1 + s_{\min} v_k)\beta^k = \beta^k. \quad \square$$

We continue with the second statement and the uniqueness of a minimizer w.r.t. its used model groups.

Lemma 5.19 (Uniqueness of the minimizer)**.** Let C be positive definite and β be a minimizer of (5.22) such that the vectors (5.26) are linearly independent. Then β is the unique minimizer in the sense that there exists no other minimizer $\widetilde{\beta}$ of (5.22) with $U_{\widetilde{\beta}} = U_\beta$.

Proof. We verify that the function J has positive definite Hessian at the minimizer β if we restrict the directions v to be small, $U_{\beta+v} = U_\beta$ and that the bias constraint is satisfied. Then J is a convex function that is strictly convex at a minimizer which must be unique. We have $\|\beta^k\|_k \neq 0$ for $k \in U_\beta$ since $\|\cdot\|_k$ is a norm for positive definite C and thus J is twice differentiable

$$\begin{aligned}\partial_{\beta_i^k \beta_j^k} J(\beta) &= \partial_{\beta_i^k \beta_j^k} \|\beta^k\|_k = \partial_{\beta_i^k \beta_j^k} \left((\beta^k)^T C^k \beta^k W^k\right)^{1/2} \\ &= W^k \frac{C_{ij}^k}{\|\beta^k\|_k} - (W^k)^2 \frac{(e_i^T C^k \beta^k)(e_j^T C^k \beta^k)}{\|\beta^k\|_k^3}.\end{aligned}$$

Notice that J as the sum of norms of β^k has a block-diagonal Hessian, thus all other second derivatives are zero. Hence, we only have to verify that the matrices

$$A^k := W^k C^k \|\beta^k\|_k^2 - (W^k)^2 (C^k \beta^k)(C^k \beta^k)^T$$

are positive definite along the directions where the bias constraint is satisfied. Since J is convex, we conclude that A^k is positive semi–definite. Now let $v^k \in \mathbb{R}^{|S^k|}$ such that $0 = (v^k)^T A^k v^k$. We denote the scalar product corresponding to the norm $\|\cdot\|_k$ with $(\cdot,\cdot)_k$ and apply the Cauchy–Schwarz inequality

$$0 = (v^k)^T A^k v^k = \|v^k\|_k^2 \|\beta^k\|_k^2 - (v^k, \beta^k)_k^2 \geq \|v^k\|_k^2 \|\beta^k\|_k^2 - (\|v^k\|_k \|\beta^k\|_k)^2 = 0.$$

The Cauchy–Schwarz inequality holds with equality if and only if v^k is a scalar multiple of β^k

$$v^k = s_k \beta^k,$$

where $s_k \in \mathbb{R}$ is arbitrary. We now demand that the directions v^k for $k \in U_\beta$ satisfy the bias constraint

$$\alpha = \sum_{k \in U_\beta} P^k(\beta^k + v^k) = \sum_{k \in U_\beta} P^k \beta^k + \sum_{k \in U_\beta} P^k v^k = \alpha + \sum_{k \in U_\beta} s_k P^k \beta^k$$

and thus we require the s_k to satisfy

$$\sum_{k \in U_\beta} s_k P^k \beta^k = 0.$$

Since the vectors $P^k \beta^k$ for $k \in U_\beta$ are linearly independent by the assumption of the theorem, the above expression is valid only if $s_k = 0$. We conclude that $v^k = 0$ and thus the Hessian of J is positive definite restricted to the linear subspace defined by the bias constraint and the constraint of only using model groups in U_β. □

The previous lemma crucially requires that the covariance matrix is positive definite, since otherwise this result does not hold. A simple counter example is $C = 0$ such that every coefficient β that satisfies the bias constraint is a an optimal solution and thus it is trivial to find a second minimizer β' with $U_{\beta'} = U_\beta$. Let us derive the final result of this section.

Theorem 5.20 (Set of minimizers)**.** Let C be positive definite. Then there exists minimizers $\beta^{[1]}, \dots, \beta^{[N]}$ of (5.22) such that the derived vectors (5.26) are linearly independent for every $\beta^{[n]}$, $n \in \{1, \dots, N\}$. Furthermore, the set of minimizers of (5.22) is the convex hull of $\beta^{[1]}, \dots, \beta^{[N]}$.

Proof. Lemma 5.19 shows that there exists at most a single minimizer β for each U_β if the set (5.26) is linear independent. Hence we have at most finitely many different minimizers with these two properties. Since Lemma 5.18 shows that every minimizer β is a convex combination of those minimizers, the result follows. □

Uniqueness under small perturbations. The previous theorem shows that the set of optimizers is up to convex combinations a discrete set and Example 5.17 shows that we have to choose the cost and covariance of the models exactly right to obtain multiple SAOBs. In this example these are the MC and MFMC estimator or any convex combination thereof. For our numerical experiments we only obtained a single unique SAOB and we now show that small perturbations of the cost or covariance lead to a unique solution. The perturbed optimization problem is given by

$$\begin{aligned} \min_{\beta} \quad & \widetilde{J}(\beta) := \sum_{k=1}^{K} \left[(\beta^k)^T C^k \beta^k (W^k + \xi_k) \right]^{1/2} \\ \text{such that} \quad & \sum_{k=1}^{K} P^k \beta^k = \alpha, \end{aligned} \tag{5.27}$$

where $\xi_1, \dots, \xi_K$ are perturbations in the cost, e.g. obtained from estimating the cost of computing the model group. Results for perturbations in the covariance, e.g. from estimation, are more challenging to examine. In any case, perturbations in the cost and covariance lead to similar results, since we may redefine the covariance of every model group to $C^k(W^k + \xi_k)$ and the cost to be equal to one without changing the minimizer or minimum of (5.27). We start with a result showing that the used model groups of the optimizers stay fixed under small perturbations in the cost.

Lemma 5.21. Let C be positive definite, $U \subseteq \{1, \dots, K\}$ and β an optimizer of the unperturbed problem (5.22) restricted to U, that is we fix $\beta^k = 0$ for $k \notin U$. Furthermore, assume that $U_\beta = U$ and that the vectors in (5.26) are linearly independent. Then for all $\varepsilon > 0$ small enough with i.i.d. perturbations $\xi_1, \dots, \xi_K \sim U(-\varepsilon, \varepsilon)$ the optimizer $\widetilde{\beta}$ for the perturbed problem (5.27) restricted to U satisfies $U_{\widetilde{\beta}} = U = U_\beta$ and the vectors in (5.26) are linearly independent with β^k replaced by $\widetilde{\beta}^k$.

Proof. We w.l.o.g. assume $\beta^1, \dots, \beta^L$ are not equal to zero and $|U_\beta| = L$. We write down the optimality conditions of (5.22) restricted to $U = \{1, \dots, L\}$ and denote the Lagrange–multiplier with $\lambda \in \mathbb{R}^L$

$$\begin{aligned} W^k C^k \beta^k - \left[(\beta^k)^T C^k \beta^k W^k \right]^{1/2} P^k \lambda = 0 \qquad & \text{for all } k \in \{1, \dots, L\}, \\ \sum_{k=1}^{L} P^k \beta^k = \alpha. & \end{aligned}$$

The optimality conditions depend continuously on W^k and thus the optimal $\beta^1, \dots, \beta^L$ also depend continuously on $W^1, \dots, W^L$. Now let $\widetilde{\beta}$ be the optimizer for the perturbed

problem restricted to U_β. Then the vectors $P^1\widetilde{\beta}^1, \dots, P^L\widetilde{\beta}^L$ are again linearly independent since the determinant is a continuous function and

$$\det(P^1\beta^1|\dots|P^L\beta^L) \neq 0.$$

due to linear independent vectors in (5.26). This also shows that the set of used models does not change under small perturbations of the costs. □

In terms of Example 5.17 Lemma 5.21 means that the MFMC estimator does not decay to a MC estimator that only uses either the model group $\{1,2\}$ or $\{1\}$ if we slightly perturb the costs.
We now assume that $U_\beta = \{1,\dots,L\}$ for the optimizer β of (5.22) such that $\Psi(m)$ is positive definite if C is positive definite. Then the optimality conditions in the formulation of the samples m (5.9) are

$$\alpha^T\Psi(m)^{-1}P^k(C^k)^{-1}R^k\Psi(m)^{-1}\alpha = W^k\lambda^{\mathbb{W}} \qquad \text{for all } k \in \{1,\dots,L\}, \tag{5.28}$$

$$\sum_{k=1}^{L} m_k W^k = \mathbb{W}^{\text{budget}}. \tag{5.29}$$

Multiplying (5.28) by m_k, summing up over $k = 1,\dots,L$ together with (5.29) and the definition of $\Psi(m)$ shows that the Lagrange–multiplier satisfies

$$\lambda^{\mathbb{W}} = \frac{\alpha^T\Psi(m)^{-1}\alpha}{\mathbb{W}^{\text{budget}}} = \frac{J(\beta)}{\mathbb{W}^{\text{budget}}}. \tag{5.30}$$

Since we only compare estimators with the same budget we examine how changes in $W^1,\dots,W^L$ changes the value of $\lambda^{\mathbb{W}}$, which we do using the implicit function theorem. We write the optimality conditions (5.28) and (5.29) as single equation

$$F(W^1,\dots,W^L,\mathbb{W}^{\text{budget}},m_1,\dots,m_L,\lambda^{\mathbb{W}}) = 0. \tag{5.31}$$

This equation is satisfied at the unperturbed optimizer. The implicit function theorem requires us to compute some derivatives, which we do now.

Lemma 5.22 (Derivatives of F)**.** Let C be positive definite and $W^1,\dots,W^L$, $\mathbb{W}^{\text{budget}}$, $m_1,\dots,m_L$, $\lambda^{\mathbb{W}}$ satisfy (5.31) with $m_k > 0$ for $k \in \{1,\dots,L\}$. Then with $W := (W^1,\dots,W^L)^T$ there holds

$$\begin{aligned}
\partial_{W^k}F &= -\lambda^{\mathbb{W}}e_k + m_k e_{L+1},\\
\partial_{\mathbb{W}^{\text{budget}}}F &= -e_{L+1},\\
\partial_{m_1,\dots,m_L,\lambda^{\mathbb{W}}}F &= \begin{pmatrix} A & -W \\ W^T & 0 \end{pmatrix}, \qquad A = (A_{k,j})_{k,j=1}^L \in \mathbb{R}^{L\times L},\\
A_{k,j} &= -2\alpha^T\Psi(m)^{-1}P^k(C^k)^{-1}R^k\Psi(m)^{-1}P^j(C^j)^{-1}R^j\Psi(m)^{-1}\alpha,\\
e_{L+1}^T(\partial_{m_1,\dots,m_L,\lambda^{\mathbb{W}}}F)^{-1} &= \frac{1}{\mathbb{W}^{\text{budget}}}(-m_1,\dots,-m_L,-2\lambda^{\mathbb{W}}).
\end{aligned}$$

Proof. The proof for the derivatives is straightforward. The matrix $-A$ is symmetric positive definite, which can be verified from

$$v^T(-A)v = 2\alpha^T\Psi(m)^{-1}\Psi(v)\Psi(m)^{-1}\Psi(v)\Psi(m)^{-1}\alpha \tag{5.32}$$

using that $\Psi(m)^{-1}$ is positive definite and $\Psi(m)$ and$\Psi(v)$ symmetric. Hence A is invertible and thus

$$(\partial_{m_1,\dots,m_L,\lambda^{\mathbb{W}}}F)^{-1} = \begin{pmatrix} A^{-1} - A^{-1}W(W^TA^{-1}W)^{-1}W^TA^{-1} & A^{-1}W(W^TA^{-1}W)^{-1} \\ -(W^TA^{-1}W)^{-1}W^TA^{-1} & (W^TA^{-1}W)^{-1} \end{pmatrix}. \tag{5.33}$$

We have $(m_1,\dots,m_L)^T = -2\lambda^{\mathbb{W}}A^{-1}W$ since

$$\begin{aligned} -A(m_1,\dots,m_L)^T &= 2\left(\sum_{k=1}^{L} m_k\alpha^T\Psi(m)^{-1}P^k(C^k)^{-1}R^k\Psi(m)^{-1}P^j(C^j)^{-1}R^j\Psi(m)^{-1}\alpha\right)_{j=1}^{L} \\ &= 2\big(\alpha^T\Psi(m)^{-1}\Psi(m)\Psi(m)^{-1}P^j(C^j)^{-1}R^j\Psi(m)^{-1}\alpha\big)_{j=1}^{L} \\ &= 2\big(\alpha^T\Psi(m)^{-1}P^j(C^j)^{-1}R^j\Psi(m)^{-1}\alpha\big)_{j=1}^{L}. \end{aligned}$$

Now use the optimality condition (5.28)

$$-A(m_1,\dots,m_L)^T = 2(W^j\lambda^{\mathbb{W}})_{j=1}^{L} = 2\lambda^{\mathbb{W}}W. \tag{5.34}$$

We use that $W^T(m_1,\dots,m_L)^T = \mathbb{W}^{\text{budget}}$ from the cost constraint (5.29) to conclude

$$(W^TA^{-1}W)^{-1} = -2\lambda^{\mathbb{W}}(W^T(-2\lambda^{\mathbb{W}}A^{-1}W))^{-1} = -2\lambda^{\mathbb{W}}(W^T(m_1,\dots,m_L)^T)^{-1} = \frac{-2\lambda^{\mathbb{W}}}{\mathbb{W}^{\text{budget}}}. \tag{5.35}$$

This now shows that the inverse in (5.33) is well defined, since (5.35) is a well–defined negative number. We now conclude the lemma with a calculation using (5.34) and (5.35)

$$\begin{aligned} e_{L+1}^T(\partial_{m_1,\dots,m_L,\lambda^{\mathbb{W}}}F)^{-1} &= (-(W^TA^{-1}W)^{-1}W^TA^{-1}, (W^TA^{-1}W)^{-1}) \\ &= \frac{1}{\mathbb{W}^{\text{budget}}}(2\lambda^{\mathbb{W}}A^{-1}W^T, -2\lambda^{\mathbb{W}}) \\ &= \frac{1}{\mathbb{W}^{\text{budget}}}(-m_1,\dots,-m_L,-2\lambda^{\mathbb{W}}). \end{aligned}$$ □

We now derive the Taylor expansion for the optimal objective value at a minimizer w.r.t. the costs of the model groups.

Lemma 5.23 (Taylor expansion)**.** Let C be positive definite and $W^1,\dots,W^L$, $\mathbb{W}^{\text{budget}}$, $m_1,\dots,m_L$, $\lambda^{\mathbb{W}}$ satisfy (5.31) with $m_k > 0$ for $k \in \{1,\dots,L\}$. Let β be the respective coefficients and $\widetilde{\beta}$ the coefficients of the perturbed problem with $U_\beta = U_{\widetilde{\beta}}$ with i.i.d. perturbations $\xi_1,\dots,\xi_K \sim U(-\varepsilon,\varepsilon)$ and $\varepsilon > 0$ sufficiently small. Then the objective function value at the minimizer satisfies the Taylor expansion

$$\widetilde{J}(\widetilde{\beta}) = J(\beta) + \frac{J(\beta)}{\mathbb{W}^{\text{budget}}}\sum_{k=1}^{L} m_k\xi_k + o(\|\xi\|). \tag{5.36}$$

Proof. We use the implicit function theorem to achieve

$$0 = F(W^1+\xi_1,\dots,W^L+\xi_L,\mathbb{W}^{\text{budget}},\widetilde{m}_1,\dots,\widetilde{m}_L,\widetilde{\lambda}^{\mathbb{W}}) = F(W+\xi,\mathbb{W}^{\text{budget}},g(W+\xi))$$

in a small neighbourhood around $(W^1,\dots,W^K,\mathbb{W}^{\text{budget}})^T \in \mathbb{R}^{L+1}$. Here $g:\mathbb{R}^{L+1}\to\mathbb{R}^{L+1}$ is a smooth function and we are only interested in the last component since (5.30) shows

$$e_{L+1}^Tg = \lambda^{\mathbb{W}} = \frac{J(\beta)}{\mathbb{W}^{\text{budget}}}.$$

The implicit function theorem and Lemma 5.22 now shows that

$$\begin{aligned}\partial_{W^k}\lambda^{\mathbb{W}} &= \partial_{W^k} e_{L+1}^T g = -e_{L+1}^T(\partial_{m_1,\dots,m_L,\lambda^{\mathbb{W}}}F)^{-1}\partial_{W^k}F \\ &= \frac{1}{\mathbb{W}^{\text{budget}}}(m_1,\dots,m_L,2\lambda^{\mathbb{W}})\big(-\lambda^{\mathbb{W}}e_k + m_k e_{L+1}\big) = \frac{\lambda^{\mathbb{W}}}{\mathbb{W}^{\text{budget}}}m_k.\end{aligned} \tag{5.37}$$

We now use a Taylor expansion, (5.37) and (5.30) to obtain the result (5.36)

$$\begin{aligned}\widetilde{J}(\widetilde{\beta}) &= \mathbb{W}^{\text{budget}}\lambda^{\mathbb{W}}(W+\xi) \\ &= \mathbb{W}^{\text{budget}}\lambda^{\mathbb{W}}(W) + \mathbb{W}^{\text{budget}}(\partial_{W^1}\lambda^{\mathbb{W}}(W),\dots,\partial_{W^L}\lambda^{\mathbb{W}}(W))(\xi_1,\dots,\xi_L)^T + o(\|\xi\|) \\ &= J(\beta) + \lambda^{\mathbb{W}}(W)\sum_{k=1}^{L} m_k\xi_k + o(\|\xi\|).\end{aligned}$$

Here $\lambda^{\mathbb{W}}(W+\xi)$ denotes the optimal Lagrange–multiplier of the perturbed and $\lambda^{\mathbb{W}}(W)$ of the unperturbed problem. □

The Taylor expansion (5.36) shows that a perturbation which increases the cost of a single model group leads to an increase in the variance proportional to the number of samples of this model group (ignoring higher order terms). The scaling factor $J(\beta)/\mathbb{W}^{\text{budget}} = \lambda^{\mathbb{W}}$ is the variance per cost ratio of the estimator at the minimizer which is multiplied by the cost change keeping the old number of samples to obtain the (linear) variance change. The result in Lemma 5.23 is thus not surprising.
We now derive the main uniqueness result.

Theorem 5.24 (Almost sure uniqueness under perturbations of W^k)**.** Let C be positive definite, $\varepsilon > 0$ sufficiently small and $\xi_1,\dots,\xi_K \sim U(-\varepsilon,\varepsilon)$ i.i.d. uniformly distributed. Then problem (5.27) has $\mathbb{P}$–almost surely a unique solution.

Proof. We write down the Taylor expansion for minimizers restricted to a general U_β, which is a straightforward generalization of (5.36)

$$\widetilde{J}(\widetilde{\beta}) = J(\beta) + \frac{J(\beta)}{\mathbb{W}^{\text{budget}}}\sum_{k\in U_\beta} m_k\xi_k + o(\|\xi\|). \tag{5.38}$$

Now assume that two minimizers $\beta^{[1]}$ and $\beta^{[2]}$ with $U_{\beta^{[1]}} \neq U_{\beta^{[2]}}$ satisfy $J(\beta^{[1]}) = J(\beta^{[2]})$ and denote the respective number of samples with $m^{[1]}$ and $m^{[2]}$. Then we use (5.38) to conclude that

$$\begin{aligned}\widetilde{J}(\widetilde{\beta}^{[1]}) &= J(\beta^{[1]}) + \frac{J(\beta^{[1]})}{\mathbb{W}^{\text{budget}}}\sum_{k\in U_{\beta^{[1]}}} m_k^{[1]}\xi_k + o(\|\xi\|) \\ &= J(\beta^{[2]}) + \frac{J(\beta^{[2]})}{\mathbb{W}^{\text{budget}}}\sum_{k\in U_{\beta^{[2]}}} m_k^{[2]}\xi_k + o(\|\xi\|) = \widetilde{J}(\widetilde{\beta}^{[2]})\end{aligned}$$

happens $\mathbb{P}$–almost never since we have a non–zero i.i.d. contribution of a random variable $m_k\xi_k$ not appearing on both sides as $U_{\beta^{[1]}} \neq U_{\beta^{[2]}}$. Hence, if $J(\beta^{[1]}) = J(\beta^{[2]})$ for the unperturbed problem we almost surely have that either $\widetilde{\beta}^{[1]}$ or $\widetilde{\beta}^{[2]}$ has a smaller objective function value than the other. We apply this idea for the minimizers $\beta^{[1]},\dots,\beta^{[N]}$ with

$U_{\beta^{[1]}}, \dots, U_{\beta^{[N]}}$ in Theorem 5.20 of the unperturbed problem. This means that we $\mathbb{P}$–almost surely obtain a unique minimizer if we restrict the problem to the model groups given as union of $U_{\beta^{[1]}}, \dots, U_{\beta^{[N]}}$.
All that remains is to verify that there are no other solutions. Let U be such that (5.22) restricted to U has the solution β with $J(\beta) > J^{\min}$, where $J^{\min}$ denotes the unrestricted minimum. Now we verify that for sufficiently small ε we still have $\widetilde{J}(\widetilde{\beta}) > \widetilde{J}^{\min}$ for the perturbed problem, where $\widetilde{\beta}$ is the respective minimizer and $\widetilde{J}^{\min}$ the unrestricted minimum of the perturbed problem. We use the continuity of J, that $\xi_1, \dots, \xi_K$ are small and that β minimizes J to verify that

$$\begin{aligned}\widetilde{J}(\widetilde{\beta}) &= \sum_{k=1}^{K} \left[(\widetilde{\beta}^k)^T C^k \widetilde{\beta}^k (W^k + \xi_k) \right]^{1/2} \geq (1-\delta) \sum_{k=1}^{K} \left[(\widetilde{\beta}^k)^T C^k \widetilde{\beta}^k W^k \right]^{1/2} \\ &= (1-\delta) J(\widetilde{\beta}) \geq (1-\delta) J(\beta)\end{aligned}$$

holds for some $\delta > 0$ that can be made arbitrarily small if we make ε sufficiently small. The result $J^{\min} \geq (1-\delta)\widetilde{J}^{\min}$ follows similarly. We thus conclude

$$\widetilde{J}(\widetilde{\beta}) \geq (1-\delta) J(\beta) = (1-\delta) \frac{J(\beta)}{J^{\min}} J^{\min} \geq (1-\delta)^2 \frac{J(\beta)}{J^{\min}} \widetilde{J}^{\min}.$$

Now use $J(\beta)/J^{\min} > 1$ and that δ is sufficiently small to obtain $\widetilde{J}(\widetilde{\beta}) > \widetilde{J}^{\min}$. □

The previous result shows that small random perturbations in W^k lead to unique solutions. For Example 5.17, where we had to choose the cost and correlation exactly right, this means that either the MC or the MFMC estimator has a smaller variance after the random perturbation.
Let us now recall the original definition of the cost per model group

$$W^k = \sum_{\ell \in S^k} w_\ell \quad \text{for all } k \in \{1, \dots, K\}, \tag{5.39}$$

where w_ℓ denotes the cost for one evaluation of Z_ℓ. It is straightforward to derive results for perturbations $\xi_1, \dots, \xi_L$ in $w_1, \dots, w_L$ instead of W^k. The perturbed problem is then

$$\begin{aligned} \min_{\beta} \quad & \widetilde{J}(\beta) := \sum_{k=1}^{K} \left[(\beta^k)^T C^k \beta^k \sum_{\ell \in S^k} (w_\ell + \xi_\ell) \right]^{1/2} \\ \text{such that} \quad & \sum_{k=1}^{K} P^k \beta^k = \alpha. \end{aligned}$$

We use the chain rule and (5.37) to compute the derivative of $\lambda^{\mathbb{W}}$ w.r.t. w_ℓ

$$\begin{aligned}\partial_{w_\ell} \lambda^{\mathbb{W}}(W) &= (\partial_{W^1, \dots, W^L} \lambda^{\mathbb{W}}(W))(\partial_{w_\ell} W) = \frac{\lambda^{\mathbb{W}}(W)}{\mathbb{W}^{\text{budget}}} (m_1, \dots, m_L) \begin{pmatrix} 1_{S^1}(\ell) \\ \vdots \\ 1_{S^L}(\ell) \end{pmatrix} \\ &= \frac{\lambda^{\mathbb{W}}(W)}{\mathbb{W}^{\text{budget}}} \sum_{\{k \,|\, \ell \in S^k\}} m_k,\end{aligned} \tag{5.40}$$

where $1_{S^j}(\ell)$ is equal to one if $\ell \in S^j$ and zero otherwise. Thus we are now proportional to the total number of evaluations of Z_ℓ instead of the number of samples of a model group. No other results change. The Taylor expansion is comparable to (5.38) and reads

$$\widetilde{J}(\widetilde{\beta}) = J(\beta) + \frac{J(\beta)}{\mathbb{W}^{\text{budget}}} \sum_{\ell=1}^{L} \left(\sum_{\{k \,|\, \ell \in S^k\}} m_k \right) \xi_\ell + o(\|\xi\|).$$

The linear change in the variance of two estimators is thus $\mathbb{P}$–almost surely the same if the number of evaluations of each model Z_ℓ is the same since we otherwise have differently scaled i.i.d. combinations of the random noise similar to the proof of Theorem 5.24. We thus obtain the following corollary, which can be proven similarly to Theorem 5.24.

Corollary 5.25 (Almost sure uniqueness under perturbations of w_ℓ)**.** Let C be positive definite, $\varepsilon > 0$ sufficiently small and $\xi_1, \dots, \xi_L \sim U(-\varepsilon, \varepsilon)$ i.i.d. uniformly distributed. Furthermore, assume that the unperturbed problem (5.22) has minimizers $\beta^{[1]}, \dots, \beta^{[N]}$ from Theorem 5.20 with respective number of samples $m^{[1]}, \dots, m^{[N]}$. Furthermore, assume that the total number of evaluations for at least one model is pairwise different, that is for all $i, j \in \{1, \dots, N\}$

$$\sum_{\{k \,|\, \ell \in S^k\}} m_k^{[i]} \neq \sum_{\{k \,|\, \ell \in S^k\}} m_k^{[j]} \qquad \text{for at least one } \ell \in \{1, \dots, L\}. \tag{5.41}$$

Then problem (5.40) has $\mathbb{P}$–almost surely a unique solution. ◇

In the setting of Example 5.17 this corollary ensures that we have almost sure uniqueness under perturbations of w_ℓ, since the MC estimator does not use the coarse model Z_1 whereas the MFMC estimator does. This also means that if we restrict the optimization to a set of model groups U such that the linear system of equations

$$\sum_{\{k \in U \,|\, \ell \in S^k\}} m_k = 0 \qquad \text{for all } \ell \in \{1, \dots, L\}$$

has the unique solution $m_k = 0$ for all $k \in U$, then (5.41) is always satisfied. As an example, assume that we optimize only over the model groups of the MLMC estimator

$$S^1 = \{1\}, S^2 = \{1, 2\}, S^3 = \{2, 3\}, \dots, S^L = \{L-1, L\}.$$

Now let $m^{[1]}$ be a minimizer and we look for a second minimizer $m^{[2]}$ such that (5.41) is not satisfied, that is

$$\sum_{\{k \in \{1,\dots,L\} \,|\, \ell \in S^k\}} m_k^{[2]} = \sum_{\{k \in \{1,\dots,L\} \,|\, \ell \in S^k\}} m_k^{[1]} \qquad \text{for all } \ell \in \{1, \dots, L\}.$$

For L this condition is

$$m_L^{[2]} = m_L^{[1]}$$

and for $L-1$ we have

$$m_L^{[2]} + m_{L-1}^{[2]} = m_L^{[1]} + m_{L-1}^{[1]}$$

and thus also $m_{L-1}^{[2]} = m_{L-1}^{[1]}$. A further recursion then shows $m^{[2]} = m^{[1]}$ and thus (5.41) is always satisfied and we $\mathbb{P}$–almost surely obtain a unique solution.
As of the writing of the thesis it is not clear whether Corollary 5.25 can be improved by removing the assumption (5.41) on the optimizer $\beta^{[1]}, \dots, \beta^{[N]}$ of the unperturbed problem. However, it is clear from Example 5.17 that we require at least three models, that is $L \geq 3$ which complicates finding a counterexample.

Chapter 6

Asymptotics of the SAOB

In this chapter we analyse the asymptotic behaviour of the SAOB for the convergence of the models to the true QoI $Z_L \to Z$ for $L \to +\infty$. The goal is to construct an estimator that is cheap in the sense that the cost does not increase too fast. Ideally, we want to recover the optimal MSE of ε^2 with costs of order ε^{-2} that is achieved with the standard MC estimator by sampling directly from Z, which is often not possible. We mainly analyse hierarchical models that exhibit high smoothness properties, which we exploit using *Richardson Extrapolation* (RE). This chapter is divided into the following sections:

- Section 6.1 is concerned with the asymptotic analysis of general linear unbiased estimators. Here we obtain a specific complexity theorem for the estimator as well as an upper bound on the number of samples. This result is a generalization of the complexity theorem of the MC and MLMC estimator, however, this result relies on the knowledge of the coefficients β^k and is thus not suitable to study the SAOB. We then proceed to derive that the SAOB is the linear unbiased estimator that has the smallest asymptotic cost. Thus the goal of the following sections is to find a good estimator to compare to.

- We introduce the RE estimators in Section 6.2, which first consists of a general repetition of RE, which we then apply for estimation. The RE estimators are constructed as a generalization of the MLMC estimator and improve the asymptotic complexity compared to the latter. In particular, we decouple the bias vector α from the variance reduction with the help of a weighted RE estimator. We finish this section with the remark that if the QoI has an analytic error expansion, then the cost for the estimation can always be reduced to the optimal ε^{-2}.

- We derive asymptotic lower bounds on the complexity of any linear unbiased estimator in Section 6.3. Clearly, if the asymptotic cost is upper bounded by ε^{-2}, then this bound is sharp. For the case of dominating costs on the fine level we show that an asymptotic expression for the lower variance bound $\mathbb{V}^{\min}$ can be used to derive a lower bound on the complexity. In particular, this allows us to derive a lower bound on the complexity for every linear unbiased estimator.

- In Section 6.4 we study some interesting academic numerical examples. We verify that the coefficients of the SAOB converge to the coefficients of the RE estimator in a particular setting. We then proceed to show that this is not always the case. In particular, the SAOB may be a strict improvement over the RE estimator in terms of variance reduction.

- In Section 6.5 we compare the estimators presented in this thesis for a smooth PDE example and numerically verify some results of this chapter. In particular, we show that under certain circumstances the SAOB and RE estimators improve over the MC estimator. We further give a possible explanation why the CV and MFMC estimators do not improve asymptotically over the MLMC estimator, even though they couple more models.

This chapter contains results from [126] and [125].

6.1 Asymptotic analysis for linear unbiased estimators

General linear unbiased estimators. We start by analysing the asymptotic complexity of a general linear unbiased estimator $\widehat{\mu}_\alpha$. The goal is to obtain a small MSE of order $\varepsilon^2 > 0$, which is the squared bias plus the variance

$$\mathrm{MSE}[\widehat{\mu}_\alpha] = \mathrm{Bias}[\widehat{\mu}_\alpha]^2 + \mathbb{V}[\widehat{\mu}_\alpha] = (\alpha^T \mu - \mathbb{E}[Z])^2 + \mathbb{V}[\widehat{\mu}_\alpha].$$

We achieve the desired accuracy if both the bias and the variance is sufficiently small

$$(\alpha^T \mu - \mathbb{E}[Z])^2 \leq \varepsilon^2/2, \tag{6.1}$$
$$\mathbb{V}[\widehat{\mu}_\alpha] \leq \varepsilon^2/2. \tag{6.2}$$

We satisfy (6.1) by choosing $\alpha = e_L$ and the level L large enough. The variance constraint (6.2) is achieved if we compute the cost of the linear unbiased estimator from (5.16) with $J(m) = \varepsilon^2/2$

$$\mathbb{W}[\widehat{\mu}_\alpha] = 2\varepsilon^{-2}\left(\sum_{k=1}^{K}\left((\beta^k)^T C^k \beta^k W^k\right)^{1/2}\right)^2. \tag{6.3}$$

Here the coefficients β^k have to satisfy the bias constraint

$$\alpha = \sum_{k=1}^{K} P^k \beta^k.$$

Importantly, the finest level L and thus the bias α and the variance depends on the MSE ε^2. Therefore, we view the estimator $\widehat{\mu}_\alpha$ as a sequence of estimators $(\widehat{\mu}_{\alpha^L})_{L=1}^\infty$ with different biases $(\alpha^L)_{L=1}^\infty$. Then clearly, also the number of model groups K, the model groups S^k and the coefficients β^k depend on L. It is sensible to require that $\alpha^L \in \mathbb{R}^L$ and that we only use the first L models, thus $\mu \in \mathbb{R}^L$. We often do not mention the extra dependence on L explicitly to keep the notation simple.
The cost in (6.3) does not include the cost for ceiling the number of samples which effects the asymptotic complexity compared to e.g. MLMC in Theorem 3.49. It is therefore desirable to decrease the rounding costs as much as possible and we achieve a useful bound if we only use L model groups $S^1, \dots, S^L$ such that $m_{L+1} = \cdots = m_K = 0$. This requirement can always be achieved according to Theorem 5.16 by changing the coefficients without increasing the variance of the estimator. Then, the ceiling costs affect at most L model groups, which yields a tractable bound. We combine this with upper bounds for the expressions $(\beta^k)^T C^k \beta^k W^k$ in (6.3) to obtain an explicit expression for the asymptotic cost for a wide range of linear unbiased estimators.

Theorem 6.1 (Asymptotic complexity for linear unbiased estimators)**.** Assume that the biases satisfy

$$|(\alpha^L)^T \mu - \mathbb{E}[Z]| \leq c 2^{-\gamma_{\mathrm{Bias}} L} \quad \text{for all } L \in \mathbb{N}. \tag{6.4}$$

Moreover, assume that for all L the coefficients β^k of $\widehat{\mu}_{\alpha^L}$ satisfy

$$\beta^k = 0 \qquad \text{for all } k \in \{L+1, \dots, K\}, \tag{6.5}$$

$$\alpha^L = \sum_{k=1}^{L} P^k \beta^k, \tag{6.6}$$

$$(\beta^k)^T C^k \beta^k \leq c 2^{-\gamma_{\text{Var}} k} \qquad \text{for all } k \in \{1, \dots, L\}, \tag{6.7}$$

$$W^k \leq c 2^{\gamma_{\text{Cost}} k} \qquad \text{for all } k \in \{1, \dots, L\}. \tag{6.8}$$

Then for all $\varepsilon \in (0, 1/e]$ there exists $L \in \mathbb{N}$ such that $\text{MSE}[\widehat{\mu}_{\alpha^L}] \leq \varepsilon^2$ with costs

$$\mathbb{W}[\widehat{\mu}_{\alpha^L}] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} + c \begin{cases} \varepsilon^{-2}, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ \varepsilon^{-2} \log(\varepsilon)^2, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-2 - \frac{\gamma_{\text{Cost}} - \gamma_{\text{Var}}}{\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}}. \end{cases} \tag{6.9}$$

Proof. The proof is a straightforward generalization of the proof of Theorem 3.31 and [31, Theorem 1], where we replace $\mathbb{V}[Z_\ell - Z_{\ell-1}]$ for the MLMC estimator with the more general $(\beta^k)^T C^k \beta^k$. For completeness, we repeat the main arguments. We choose L in (6.4) such that (6.1) is satisfied

$$L \geq -\frac{\log_2(\varepsilon)}{\gamma_{\text{Bias}}} + c. \tag{6.10}$$

We apply Theorem 5.13 and compute the cost according to (6.3) to obtain (6.2) such that the estimator has a MSE of ε^2. We use the variance reduction (6.7) and geometric cost increase (6.8) for the L active model groups due to (6.5)

$$\mathbb{W}[\widehat{\mu}_\alpha] = 2\varepsilon^{-2} \left(\sum_{k=1}^{L} \left((\beta^k)^T C^k \beta^k W^k \right)^{1/2} \right)^2 \leq c\varepsilon^{-2} \left(\sum_{k=1}^{L} 2^{\frac{k}{2}(\gamma_{\text{Cost}} - \gamma_{\text{Var}})} \right)^2.$$

For brevity we now only verify the case $\gamma_{\text{Var}} > \gamma_{\text{Cost}}$. Here the geometric sum with negative exponent is bounded independently of L and thus of ε. Hence the cost excluding rounding is

$$\mathbb{W}[\widehat{\mu}_\alpha] \leq c\varepsilon^{-2}.$$

Finally, the cost for ceiling is upper bounded by a single evaluation of the model group W^L since we assumed a geometrical cost increase (6.8). This adds costs of order

$$\mathbb{W}[Z_L] \leq W^L \leq c 2^{\gamma_{\text{Cost}} L} \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}$$

and shows (6.9) for $\gamma_{\text{Var}} > \gamma_{\text{Cost}}$. The other cases $\gamma_{\text{Var}} = \gamma_{\text{Cost}}$ and $\gamma_{\text{Var}} < \gamma_{\text{Cost}}$ follow similarly. □

We derive an asymptotic upper bound for the number of samples on every level. This result is interesting for practical considerations, since we often have to compute the sample covariance matrix. If the number of high fidelity models tends to infinity and the cost increase is geometric, then this allows us to use more and more pilot samples for the sample covariance matrix without effecting the total cost of the SAOB too much.

Corollary 6.2 (Asymptotic upper bound on the number of samples)**.**
Let the assumptions of Theorem 6.1 be true and additionally assume that the model group cost is lower bounded

$$W^k \geq c2^{\gamma_{\text{Cost}}k} \qquad \text{for all } k \in \{1, \dots, L\}. \tag{6.11}$$

Then for all $\varepsilon \in (0, 1/e]$ the number of samples on level $k \in \{1, \dots, L\}$ to achieve $\text{MSE}[\widehat{\mu}_{\alpha^L}] \leq \varepsilon^2$ is upper bounded by

$$m_k \leq 1 + c\varepsilon^{-2+\frac{\gamma_{\text{Var}}+\gamma_{\text{Cost}}}{2\gamma_{\text{Bias}}}\frac{k}{L}} \begin{cases} 1, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ |\log(\varepsilon)|, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{2\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}}. \end{cases} \tag{6.12}$$

Proof. We use (5.15), (5.16) and parts of the proof of Theorem 6.1 to conclude

$$\begin{aligned} m_k &= \varepsilon^{-2}\left[\sum_{j=1}^{K}\left((\beta^j)^T C^j \beta^j W^j\right)^{1/2}\right]\left(\frac{(\beta^k)^T C^k \beta^k}{W^k}\right)^{1/2} \\ &\leq \varepsilon^{-1}\sqrt{\phi(\varepsilon)}\left(\frac{(\beta^k)^T C^k \beta^k}{W^k}\right)^{1/2}, \end{aligned}$$

where $\phi(\varepsilon)$ is the asymptotic cost in (6.9) with the costs for rounding excluded

$$\phi(\varepsilon) := \varepsilon^{-2}\begin{cases} 1, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ \log(\varepsilon)^2, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}}. \end{cases}$$

We use (6.7), (6.11) and insert the expression for L in (6.10) to obtain

$$\left(\frac{(\beta^k)^T C^k \beta^k}{W^k}\right)^{1/2} \leq c2^{-\frac{\gamma_{\text{Var}}+\gamma_{\text{Cost}}}{2}k} = c2^{-\frac{\gamma_{\text{Var}}+\gamma_{\text{Cost}}}{2}L\frac{k}{L}} \leq c\varepsilon^{\frac{\gamma_{\text{Var}}+\gamma_{\text{Cost}}}{2\gamma_{\text{Bias}}}\frac{k}{L}}.$$

We combine the results of this proof and ceil the number of samples to conclude (6.12). □

Upper bounds for SAOB. It is straightforward to verify that for $\alpha^L = e_L$ and coefficients β^k of the MC or MLMC estimator from Section 4.4 we obtain the corresponding asymptotic results Theorem 3.11 and Theorem 3.49. It is however not straightforward to analyse the SAOB, since the coefficients β^k are chosen implicitly by solving (5.22), where even uniqueness of the coefficients is not guaranteed. Therefore, Theorem 6.1 cannot be used to analyse the SAOB. For the same reason, this theorem cannot be used to analyse the ACV and MFMC estimators. However, since the SAOB achieves the smallest variance its costs are not larger than any of the other estimators.

Theorem 6.3 (Asymptotic optimality of SAOB)**.** Let $\widehat{\mu}_\alpha$ be a linear unbiased estimator that uses m_k i.i.d. samples of S^k for all $k \in \{1, \dots, K\}$, where the samples are also independent across model groups. Furthermore, assume that

$$\text{MSE}[\widehat{\mu}_\alpha] \leq \varepsilon^2 \quad \text{with } \mathbb{W}[\widehat{\mu}_\alpha] \leq \phi(\varepsilon).$$

Then the SAOB achieves the same accuracy with the smaller or equal costs

$$\mathrm{MSE}[\widehat{\mu}_\alpha^{\mathrm{SAOB}}] \leq \varepsilon^2 \quad \text{with } \mathbb{W}\big[\widehat{\mu}_\alpha^{\mathrm{SAOB}}\big] \leq \phi(\varepsilon). \tag{6.13}$$

In particular, if every estimator of the sequence $(\widehat{\mu}_{\alpha^L})_{L=1}^\infty$ satisfies the assumptions of this theorem, then the asymptotic cost of the sequence $(\widehat{\mu}_{\alpha^L}^{\mathrm{SAOB}})_{L=1}^\infty$ is never larger compared to the cost of $(\widehat{\mu}_{\alpha^L})_{L=1}^\infty$.

Proof. We use a straightforward bias-variance decomposition. The SAOB $\widehat{\mu}_\alpha^{\mathrm{SAOB}}$ has by definition the same bias as $\widehat{\mu}_\alpha$, thus the only difference is their variances. Since the SAOB has equal or smaller variance by Theorem 5.2, the statement (6.13) holds. □

It is straightforward to generalize the previous theorem to other kinds of estimators where some model groups S^k are not used or only a fixed amount of time. We write this down for the special case of the SAOB κ.

Corollary 6.4 (Asymptotic optimality of SAOB κ)**.** Let $\widehat{\mu}_\alpha$ be a linear unbiased estimator that uses m_k i.i.d. samples of S^k for all $k \in \{1, \dots, K\}$ with $m_k = 0$ if $|S^k| > \kappa$, where the samples are also independent across model groups. Furthermore, assume that

$$\mathrm{MSE}[\widehat{\mu}_\alpha] \leq \varepsilon^2 \quad \text{with } \mathbb{W}[\widehat{\mu}_\alpha] \leq \phi(\varepsilon).$$

Then the SAOB κ achieves the same accuracy with the smaller or equal costs

$$\mathrm{MSE}[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa}] \leq \varepsilon^2 \quad \text{with } \mathbb{W}\big[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa}\big] \leq \phi(\varepsilon).$$

In particular, if every estimator of the sequence $(\widehat{\mu}_{\alpha^L})_{L=1}^\infty$ satisfies the assumptions of this theorem, then the asymptotic cost of the sequence $(\widehat{\mu}_{\alpha^L}^{\mathrm{SAOB}\,\kappa})_{L=1}^\infty$ is never worse compared to $(\widehat{\mu}_{\alpha^L})_{L=1}^\infty$. ◇

We informally summarize Theorem 6.3 and Corollary 6.4.

- The SAOB is asymptotically optimal in the class of linear unbiased estimators.
- The SAOB κ is asymptotically optimal in the class of linear unbiased estimators that couple at most κ models.

Here the phrase "in the class of linear unbiased estimators" means that we have independent evaluations for samples in a model group and across different model groups. This shows that the SAOBs have decreasing asymptotic cost with increasing coupling κ.

Corollary 6.5 (Cost ordering for SAOB and SAOB κ)**.** Assume that the SAOB κ achieves

$$\mathrm{MSE}[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa}] \leq \varepsilon^2 \quad \text{with } \mathbb{W}\big[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa}\big] \leq \phi(\varepsilon).$$

Then for all $\kappa' \geq \kappa$

$$\mathrm{MSE}[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa'}] \leq \varepsilon^2 \quad \text{with } \mathbb{W}\Big[\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa'}\Big] \leq \phi(\varepsilon). \tag{6.14}$$

In particular, we can replace $\widehat{\mu}_\alpha^{\mathrm{SAOB}\,\kappa'}$ with $\widehat{\mu}_\alpha^{\mathrm{SAOB}}$ in (6.14). ◇

We now use Theorem 6.3 for upper complexity bounds of the SAOB: Find a sequence of comparison estimators $(\widehat{\mu}_{\alpha^L})_{L=1}^{\infty}$ where the asymptotic complexity $\phi(\varepsilon)$ is explicitly known and use this as bound for SAOB. Similarly, for the SAOB κ we apply Corollary 6.4, where we additionally have to ensure that the comparison estimator couples at most κ models. Furthermore, this allows us to choose the biases $(\alpha^L)_{L=1}^{\infty}$ from the comparison estimator. This is important, since up until now we assumed that the bias for the SAOB is given a priori.
The question is now of course, how to construct the comparison estimators such that the cost bound $\phi(\varepsilon)$ is as small as possible. Observe that the coefficients of the SAOB solve the cost minimization problem (5.22)

$$\begin{aligned} \min_{\beta} J(\beta) &:= \sum_{k=1}^{K} \left((\beta^k)^T C^k \beta^k W^k\right)^{1/2} \\ \text{such that} \qquad \alpha^L &= \sum_{k=1}^{K} P^k \beta^k. \end{aligned} \tag{6.15}$$

Therefore we use the key Theorem 6.3 and the following strategy to obtain a sequence of comparison estimators and thus an upper bound on the complexity of the SAOB:

- Suitably choose the sequence of biases $(\alpha^L)_{L=1}^{\infty}$ such that the bias convergence rate γ_{Bias} in (6.4) is large.
- Choose explicitly known coefficients β^k such that the variance $(\beta^k)^T C^k \beta^k$ in (6.7) is asymptotically small and γ_{Var} is explicitly known. Additionally, ensure that the bias constraint (6.6) as well as the geometric cost increase (6.8) with known γ_{Cost} is satisfied. Theorem 5.16 ensures that (6.5) is satisfied. Then also J in (6.15) is small, which is the cost of the estimator $\widehat{\mu}_{\alpha^L}$ with coefficients β^k and optimal (fractional) sample allocation.
- We use Theorem 6.1 to obtain the explicit asymptotic bound $\phi(\varepsilon)$ in (6.9) for the sequence of estimators $(\widehat{\mu}_{\alpha^L})_{L=1}^{\infty}$.
- Apply Theorem 6.3 to show that the SAOB has costs smaller or equal to $(\widehat{\mu}_{\alpha^L})_{L=1}^{\infty}$. Hence the explicit bound $\phi(\varepsilon)$ also holds for the SAOB.

The above approach can be generalized for the SAOB κ by using model groups S^k with coupling κ and thus $\beta^k = 0$ if $|S^k| > \kappa$. We now apply the outlined procedure for both the MC and the MLMC estimators.

Example 6.6 (MC estimator)**.** We define $\alpha^L := e_L$ and $\beta^1 := 1$ with $S^1 := \{L\}$ and zero for all other coefficients. Then we assume that (6.4) is satisfied for some rate $\gamma_{\text{Bias}} > 0$ and that (6.8) is satisfied for some $\gamma_{\text{Cost}} > 0$. Finally, if the lower bound $\mathbb{V}[Z_L] \geq c > 0$ holds independently of L, then the only sensible choice is to use $\gamma_{\text{Var}} = 0$ in (6.7). Therefore $\gamma_{\text{Var}} < \gamma_{\text{Cost}}$ in (6.9) holds and the cost to achieve a MSE of ε^2 is

$$\mathbb{W}\left[\widehat{\mu}_L^{\text{SAOB}}\right] \leq \mathbb{W}\left[\widehat{\mu}_L^{\text{SAOB}\,\kappa}\right] \leq \mathbb{W}\left[\widehat{\mu}_L^{\text{MC}}\right] \leq c\varepsilon^{-2-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{\gamma_{\text{Bias}}}} = c\varepsilon^{-2-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}.$$

The result for the MC estimator coincides with the well–known asymptotic complexity Theorem 3.11. It is straightforward to extend this result to the MC estimator that uses a different bias by redefining the QoI

$$\widehat{\mu}_{\alpha^L}^{\text{MC}} := \frac{1}{m_1}\sum_{i=1}^{m_1} \widetilde{Z}_L^i = \frac{1}{m_1}\sum_{i=1}^{m_1}\left(\sum_{\ell=1}^{L} \alpha_\ell^L Z_\ell^i\right),$$

where the bias is now $(\alpha^L)^T\mu$ with a potentially improved rate γ_{Bias}. ◇

Corollary 6.7 (MC complexity bound for SAOB and SAOB κ)**.** Assume that the rates γ_{Bias} and γ_{Cost} are positive and that for all $L \in \mathbb{N}$

$$|(\alpha^L)^T\mu - \mathbb{E}[Z]| \leq c2^{-\gamma_{\text{Bias}}L},$$
$$w_L \leq c2^{\gamma_{\text{Cost}}L}.$$

Then for all $\varepsilon \in (0, 1/e]$ the SAOB and SAOB κ with $\kappa \in \mathbb{N}$ achieve a MSE of ε^2 with costs bounded by

$$\mathbb{W}\big[\widehat{\mu}^{\text{SAOB}}_{\alpha^L}\big] \leq \mathbb{W}\big[\widehat{\mu}^{\text{SAOB}\,\kappa}_{\alpha^L}\big] \leq c\varepsilon^{-2-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}.$$ ◇

The previous corollary gives a weak and often non–tight bound for the cost of the SAOB. However, we have no assumption on the variance reduction rate γ_{Var}. Under additional assumptions the MLMC estimator achieves a better asymptotic complexity which implies a better complexity bound for the SAOB.

Example 6.8 (MLMC estimator)**.** Similarly to the MC estimator we choose $\alpha_L := e_L$ and the standard setting for MLMC

$$\begin{aligned} S^1 &:= \{1\}, & \beta^1 &:= 1, \\ S^2 &:= \{1, 2\}, & \beta^2 &:= (-1, 1)^T, \\ &\vdots & &\vdots \\ S^L &:= \{L-1, L\}, & \beta^L &:= (-1, 1)^T. \end{aligned}$$

This is a linear unbiased estimator for μ_L. Crucially, the only difference compared to the MC estimator is that

$$(\beta^\ell)^T C^\ell \beta^\ell = \mathbb{V}[Z_\ell - Z_{\ell-1}] \leq c2^{-\gamma_{\text{Var}}\ell}$$

has often a variance reduction rate $\gamma_{\text{Var}} > 0$. Therefore, we conclude (6.9) for the MLMC estimator

$$\mathbb{W}\big[\widehat{\mu}^{\text{SAOB}}_L\big] \leq \mathbb{W}\big[\widehat{\mu}^{\text{SAOB}\,\kappa}_L\big] \leq \mathbb{W}\big[\widehat{\mu}^{\text{MLMC}}_L\big] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} + c\begin{cases} \varepsilon^{-2}, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}}. \end{cases}$$

We require that $\kappa \geq 2$ for SAOB κ since the MLMC estimator couples two models. The bound can also directly be obtained from Theorem 3.49. ◇

Corollary 6.9 (MLMC complexity bound for SAOB and SAOB κ)**.** Let the rates γ_{Bias}, γ_{Var} and γ_{Cost} be non–negative and such that for all $L \in \mathbb{N}$

$$|\mu_L - \mathbb{E}[Z]| \leq c2^{-\gamma_{\text{Bias}}L},$$
$$\mathbb{V}[Z_L - Z_{L-1}] \leq c2^{-\gamma_{\text{Var}}L},$$
$$w_L \leq c2^{\gamma_{\text{Cost}}L}.$$

Then for all $\varepsilon \in (0, 1/e]$ the SAOB and SAOB κ with $\kappa \in \{2, 3, \dots\}$ achieve a MSE of ε^2 with costs bounded by

$$\mathbb{W}\big[\widehat{\mu}^{\text{SAOB}}_L\big] \leq \mathbb{W}\big[\widehat{\mu}^{\text{SAOB}\,\kappa}_L\big] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} + c\begin{cases} \varepsilon^{-2}, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}}. \end{cases}$$ ◇

We now improve the two previous examples by increasing the bias and the variance reduction rate using RE.

6.2 Richardson Extrapolation Estimator

Standard RE. Indeed, under stronger assumptions on the models we are able to use RE named after Richardson [117]. RE is a well–known technique to improve the accuracy of numerical approximations [22]. It is used for quadrature [120], for ordinary differential equations [23] and for stochastic ordinary differential equations [104, 133]. The technique was also applied to PDEs in [5, 11, 113]. This technique was already used in the original MLMC paper [56] to improve the bias. The authors of [89, 97] also considered RE to obtain an even smaller bias. Therefore, the results derived in this section concerning the mean and the bias rate γ_{Bias} are known or only slight variations of these results. However, we expand RE to also achieve an improved variance reduction rate γ_{Var} and we construct a weighted RE estimators such that we are able to choose both γ_{Bias} and γ_{Var} independently. We achieve this by separately applying RE for the mean and pathwise for the variance. We now write down the two basic model assumptions. The first assumption concerns the mean, which we will later use improve the bias. The second assumption concerns the realizations and will be used to improve the variance reduction rate.

Assumption 6.10 (Mean expansion)**.** There exists $q_{\text{mean}} \in \mathbb{N}$ and $0 = \gamma_1 < \cdots < \gamma_{q_{\text{mean}}}$ such that for all $\ell \in \mathbb{N}$

$$\mathbb{E}[Z_\ell] = \mathbb{E}[Z] + \sum_{j=2}^{q_{\text{mean}}-1} c_j 2^{-\gamma_j \ell} + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}} \ell}), \tag{6.16}$$

where $\mathcal{O}(2^{-\gamma_{q_{\text{mean}}} \ell})$ is meant in the sense of $\ell \to +\infty$. ◇

Assumption 6.11 (Pathwise expansion)**.** There exists $q_{\text{path}} \in \mathbb{N}$ and $0 = \gamma_1 < \cdots < \gamma_{q_{\text{path}}}$ such that for all $\ell \in \mathbb{N}$ and $\mathbb{P}$–almost surely

$$Z_\ell(\omega) = Z(\omega) + \sum_{j=2}^{q_{\text{path}}-1} c_j(\omega) 2^{-\gamma_j \ell} + \mathcal{O}(2^{-\gamma_{q_{\text{path}}} \ell}), \tag{6.17}$$

where $\mathcal{O}(2^{-\gamma_{q_{\text{path}}} \ell})$ is meant in the L^2-sense for $\ell \to +\infty$

$$\mathbb{E}\left[\left(Z_\ell - Z - \sum_{j=2}^{q_{\text{path}}-1} c_j 2^{-\gamma_j \ell}\right)^2\right] \leq c 2^{-2\gamma_{q_{\text{path}}} \ell} \quad \text{for all } \ell \in \mathbb{N}.$$

The random variables $c_2, \ldots, c_{q_{\text{path}}-1}$ as well as the random remainder have finite second moments. ◇

It is obvious that Assumption 6.11 implies Assumption 6.10 with $q_{\text{mean}} \geq q_{\text{path}}$ by taking the mean and using that $c_2, \ldots, c_{q_{\text{path}}-1}$ and the remainder have finite second moments. However, the converse is not true.

Example 6.12 (Assumption 6.10 $\not\Rightarrow$ Assumption 6.11)**.** Let Z_ℓ satisfy (6.17) with a remainder that is not of the form $2^{-\gamma\ell}$

$$Z_\ell(\omega) := Z(\omega) + \sum_{j=2}^{q_{\text{path}}-1} c_j(\omega)2^{-\gamma_j\ell} + c_{q_{\text{path}}}(\omega)\frac{2^{-\gamma_{q_{\text{path}}-1}\ell}}{\log(\ell)}.$$

For fixed $\delta > 0$ observe that

$$\frac{2^{-\gamma_{q_{\text{path}}-1}\ell}}{\log(\ell)} > c2^{-(\gamma_{q_{\text{path}}-1}+\delta)\ell} \quad \text{for } \ell \text{ sufficiently large}$$

and thus Assumption 6.11 holds with at most $\gamma_{q_{\text{path}}-1}$. On the other hand, if $\mathbb{E}\big[c_{q_{\text{path}}}\big] = 0$, then Assumption 6.10 holds with a zero remainder in (6.16). Thus any $q_{\text{mean}} \geq q_{\text{path}}$ is a valid choice and in particular strict inequality is possible. ◇

The intuitive explanation of the previous example is that a non–smooth contribution is averaged out to a smooth result. In this case, the mean $\mathbb{E}$ acts as smoothing of the expansion, which allows us to use more terms. RE linearly combines models to achieve a higher order of approximation. We define the RE vectors for $q \leq L$ and $k \in \{1, \dots, L\}$

$$v^{k,q} := \begin{cases} 0, & \text{if } k = 0, \\ e_1, & \text{if } k = 1, \\ (2^{\gamma_k} Dv^{k-1,q} - v^{k-1,q})/(2^{\gamma_k} - 1), & \text{if } 1 < k < q, \\ Dv^{k-1,q} & \text{if } k \geq q. \end{cases} \in \mathbb{R}^L. \tag{6.18}$$

Here the matrix D shifts a vector in the following sense

$$D := \begin{pmatrix} 0 & 0 \\ I_{L-1,L-1} & 0 \end{pmatrix} \in \mathbb{R}^{L\times L}.$$

We remark that $v_\ell^{k,q} = 0$ if $\ell > k$ and thus we view $v^{k,q} \in \mathbb{R}^L$ for some L sufficiently large. The specific linear combination (6.18) removes the terms of order 2^{γ_j} from (6.16).

Lemma 6.13 (Mean RE linear combination)**.** Let Assumption 6.10 be true. Then the linear combination $\sum_{\ell=1}^k v_\ell^{k,q} Z_{\ell+\ell_0}$ with $k \in \mathbb{N}$ and starting level $\ell_0 \in \mathbb{N}_0$ satisfies

$$\mathbb{E}\left[\sum_{\ell=1}^k v_\ell^{k,q} Z_{\ell+\ell_0}\right] = \mathbb{E}[Z] + \sum_{j=k+1}^{q_{\text{mean}}-1} c_j^k 2^{-\gamma_j\ell_0} + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}\ell_0}), \tag{6.19}$$

where c_j^k are suitable constants. Importantly, if $k+1 > q_{\text{mean}} - 1$ then the sum disappears and only $\mathbb{E}[Z] + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}\ell_0})$ remains.

Proof. The proof is a well–known result for RE. The basic idea is to use the (6.18) such that the terms of lower order cancel. We use induction over k to prove the desired statement. For $k = 1$ we have $v^{1,q} = e_1$ and thus (6.19) trivially holds from (6.16) of Assumption 6.10. Now let $1 < k < q$, use the properties of the shift matrix and $v_k^{k-1,q} = 0$ to conclude

$$\begin{aligned} \mathbb{E}\left[\sum_{\ell=1}^k v_\ell^{k,q} Z_{\ell+\ell_0}\right] &= \sum_{\ell=1}^k \frac{2^{\gamma_k}(Dv^{k-1,q})_\ell - v_\ell^{k-1,q}}{2^{\gamma_k} - 1}\mathbb{E}[Z_{\ell+\ell_0}] \\ &= \frac{2^{\gamma_k}}{2^{\gamma_k} - 1}\sum_{\ell=1}^{k-1} v_\ell^{k-1,q}\mathbb{E}[Z_{\ell+\ell_0+1}] - \frac{1}{2^{\gamma_k} - 1}\sum_{\ell=1}^{k-1} v_\ell^{k-1,q}\mathbb{E}[Z_{\ell+\ell_0}]. \end{aligned}$$

We use the induction hypothesis with $k-1$, ℓ_0+1 for the first and ℓ_0 for the second summand

$$\begin{aligned}
\mathbb{E}\left[\sum_{\ell=1}^{k} v_\ell^{k,q} Z_{\ell+\ell_0}\right] &= \frac{2^{\gamma_k}}{2^{\gamma_k}-1}\left(\mathbb{E}[Z] + \sum_{j=k}^{q_{\text{mean}}-1} c_j^{k-1} 2^{-\gamma_j(\ell_0+1)} + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}(\ell_0+1)})\right) \\
&\quad - \frac{1}{2^{\gamma_k}-1}\left(\mathbb{E}[Z] + \sum_{j=k}^{q_{\text{mean}}-1} c_j^{k-1} 2^{-\gamma_j \ell_0} + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}\ell_0})\right) \\
&= \mathbb{E}[Z] + \sum_{j=k}^{q_{\text{mean}}-1}\left(\frac{2^{\gamma_k}}{2^{\gamma_k}-1} c_j^{k-1} 2^{-\gamma_j(\ell_0+1)} - \frac{1}{2^{\gamma_k}-1} c_j^{k-1} 2^{-\gamma_j\ell_0}\right) \\
&\quad + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}\ell_0}) \\
&= \mathbb{E}[Z] + \sum_{j=k}^{q_{\text{mean}}-1} \frac{2^{\gamma_k}2^{-\gamma_j}-1}{2^{\gamma_k}-1} c_j^{k-1} 2^{-\gamma_j\ell_0} + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}\ell_0}).
\end{aligned}$$

We introduce the constants $c_j^0 := c_j$ and

$$c_j^k := \frac{2^{\gamma_k}2^{-\gamma_j}-1}{2^{\gamma_k}-1} c_j^{k-1}.$$

Crucially, we have that $c_k^k = 0$ and thus the order $2^{-\gamma_k\ell_0}$ disappears from the expansion showing (6.19). Finally, for $k \geq q$ the RE vector is constructed from the shift matrix

$$\mathbb{E}\left[\sum_{\ell=1}^{k} v_\ell^{k,q} Z_{\ell+\ell_0}\right] = \mathbb{E}\left[\sum_{\ell=1}^{k} (Dv^{k-1,q})_\ell Z_{\ell+\ell_0}\right] = \mathbb{E}\left[\sum_{\ell=1}^{k-1} v_\ell^{k-1,q} Z_{\ell+\ell_0+1}\right].$$

Hence, the result also follows from induction with $k-1=q$. □

We now also state essentially the same result for the case of the pathwise expansion. The proof is identical, except that instead of deterministic coefficients we have realizations of random variables.

Lemma 6.14 (Pathwise RE linear combination)**.** Let Assumption 6.11 be true. Then the linear combination $\sum_{\ell=1}^{k} v_\ell^{k,q} Z_{\ell+\ell_0}$ with $k \in \mathbb{N}$ and starting level $\ell_0 \in \mathbb{N}_0$ satisfies

$$\sum_{\ell=1}^{k} v_\ell^{k,q} Z_{\ell+\ell_0} = Z + \sum_{j=k+1}^{q_{\text{path}}-1} c_j^k 2^{-\gamma_j\ell_0} + \mathcal{O}(2^{-\gamma_{q_{\text{path}}}\ell_0}), \tag{6.20}$$

where c_j^k are suitable random variables with finite second moment. The remainder also has finite second moment. Importantly, if $k+1 > q_{\text{path}}-1$ then the sum disappears and only $Z + \mathcal{O}(2^{-\gamma_{q_{\text{path}}}\ell_0})$ remains.

Proof. The proof is identical to Lemma 6.13. We write down how the random variables c_j^k are defined

$$c_j^k := \frac{2^{\gamma_k}2^{-\gamma_j}-1}{2^{\gamma_k}-1} c_j^{k-1},$$

with the start of the recursion $c_j^0 := c_j$. These random variables and the remainder have finite second moments as linear combinations of random variables with finite second moments. □

Before we continue we want to derive an explicit expression for $v^{q-1,q}$, which can then be used to derive all other RE vectors.

Lemma 6.15 (Explicit expressions for RE vectors)**.** For all $q \in \{2,3,4,\dots\}$ the RE vector satisfies

$$v^{q-1,q} = \frac{1}{\prod_{j=2}^{q-1}(2^{\gamma_j}-1)} \begin{pmatrix} (-1)^q \\ (-1)^{q-1}\sum_{2\le i_1\le q-1} 2^{\gamma_{i_1}} \\ (-1)^{q-2}\sum_{2\le i_1<i_2\le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}} \\ \vdots \\ \sum_{2\le i_1<i_2<\dots<i_{q-3}\le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-3}}} \\ -\sum_{2\le i_1<i_2<\dots<i_{q-2}\le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-2}}} \\ \sum_{2\le i_1<i_2<\dots<i_{q-1}\le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-1}}} \end{pmatrix} \in \mathbb{R}^{q-1}. \tag{6.21}$$

Now we view $v^{q-1,q} \in \mathbb{R}^L$ with suitable zero extension. Then $v^{\ell,q} = v^{\ell,\ell+1}$ for all $\ell \in \{1,\dots,q-1\}$ and $v^{\ell,q} = D^{\ell-q+1}v^{q-1,q}$ for $\ell \in \{q,\dots,L\}$.

Proof. We use induction over q. For $q=2$ we have $v^{1,2} = e_1$ by its definition (6.18). For $q>2$ we again use the definition (6.18), $v^{q-2,q}_{\{1,\dots,q-2\}} = v^{q-2,q-1}$ and the induction hypothesis

$$\begin{aligned} v^{q-1,q} &= \frac{1}{2^{\gamma_{q-1}}-1}(2^{\gamma_{q-1}}Dv^{q-2,q} - v^{q-2,q}) \\ &= \frac{1}{(2^{\gamma_{q-1}}-1)\prod_{j=2}^{q-2}(2^{\gamma_j}-1)} \begin{pmatrix} 0 \\ 2^{\gamma_{q-1}}(-1)^{q-1} \\ 2^{\gamma_{q-1}}(-1)^{q-2}\sum_{2\le i_1\le q-2} 2^{\gamma_{i_1}} \\ 2^{\gamma_{q-1}}(-1)^{q-3}\sum_{2\le i_1<i_2\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}} \\ \vdots \\ 2^{\gamma_{q-1}}\sum_{2\le i_1<i_2<\dots<i_{q-4}\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-4}}} \\ -2^{\gamma_{q-1}}\sum_{2\le i_1<i_2<\dots<i_{q-3}\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-3}}} \\ 2^{\gamma_{q-1}}\sum_{2\le i_1<i_2<\dots<i_{q-2}\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-2}}} \end{pmatrix} \\ &\quad - \frac{1}{(2^{\gamma_{q-1}}-1)\prod_{j=2}^{q-2}(2^{\gamma_j}-1)} \begin{pmatrix} (-1)^{q-1} \\ (-1)^{q-2}\sum_{2\le i_1\le q-2} 2^{\gamma_{i_1}} \\ (-1)^{q-3}\sum_{2\le i_1<i_2\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}} \\ \vdots \\ \sum_{2\le i_1<i_2<\dots<i_{q-3}\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-4}}} \\ -\sum_{2\le i_1<i_2<\dots<i_{q-2}\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-3}}} \\ \sum_{2\le i_1<i_2<\dots<i_{q-1}\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\dots+\gamma_{i_{q-2}}} \\ 0 \end{pmatrix}. \end{aligned}$$

The denominator has the correct value and we further conclude that all entries of $v^{q-1,q}$ have the correct sign. We now show that the absolute value of the vector also has the correct value, i.e. for $v_4^{q-1,q}$ we have to show

$$\sum_{2\le i_1<i_2\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\gamma_{q-1}} + \sum_{2\le i_1<i_2<i_3\le q-2} 2^{\gamma_{i_1}+\gamma_{i_2}+\gamma_{i_3}} = \sum_{2\le i_1<i_2<i_3\le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\gamma_{i_3}}.$$

This is however, a combinatorial argument. All subsets of $\{2,\dots,q-1\}$ of size n can be written as union of sets

$$\begin{aligned} &\{I\cup\{q-1\} \,|\, I\subseteq\{2,\dots,q-2\}, |I| = n-1\} \cup \{I\subseteq\{2,\dots,q-2\} \,|\, |I| = n\} \\ &= \{I\subseteq\{2,\dots,q-1\} \,|\, |I| = n\}. \end{aligned}$$

The result (6.21) then follows. The remaining statement of this lemma is straightforward to verify from definition (6.18). □

Let us give some brief examples for the expression (6.21).

Example 6.16 (RE vectors)**.** We use the recursion (6.18).

$$
\begin{aligned}
v^{1,2} &= e_1,\\
v^{2,3} &= \frac{1}{2^{\gamma_2}-1}\begin{pmatrix} 0 & -1 \\ 2^{\gamma_2} & -0 \end{pmatrix} = \frac{1}{2^{\gamma_2}-1}\begin{pmatrix} -1 \\ 2^{\gamma_2} \end{pmatrix},\\
v^{3,4} &= \frac{1}{(2^{\gamma_2}-1)(2^{\gamma_3}-1)}\begin{pmatrix} 0 & -(-1) \\ 2^{\gamma_3}(-1) & -2^{\gamma_2} \\ 2^{\gamma_3}2^{\gamma_2} & -0 \end{pmatrix} = \frac{1}{(2^{\gamma_2}-1)(2^{\gamma_3}-1)}\begin{pmatrix} 1 \\ -(2^{\gamma_3}+2^{\gamma_2}) \\ 2^{\gamma_3+\gamma_2} \end{pmatrix},\\
v^{4,5} &= \frac{1}{(2^{\gamma_2}-1)(2^{\gamma_3}-1)(2^{\gamma_4}-1)}\begin{pmatrix} 0 & -1(1) \\ 2^{\gamma_4} & -(-(2^{\gamma_3}+2^{\gamma_2})) \\ 2^{\gamma_4}(-(2^{\gamma_3}+2^{\gamma_2})) & -(2^{\gamma_3+\gamma_2}) \\ 2^{\gamma_4}2^{\gamma_3+\gamma_2} & -0 \end{pmatrix}\\
&= \frac{1}{(2^{\gamma_2}-1)(2^{\gamma_3}-1)(2^{\gamma_4}-1)}\begin{pmatrix} -1 \\ 2^{\gamma_4}+2^{\gamma_3}+2^{\gamma_2} \\ -(2^{\gamma_4+\gamma_3}+2^{\gamma_4+\gamma_2}+2^{\gamma_3+\gamma_2}) \\ 2^{\gamma_4+\gamma_3+\gamma_2} \end{pmatrix}.
\end{aligned}
$$

These expressions coincide with (6.21). ◇

RE estimator. We use the same idea underlying MLMC to obtain an estimator with improved variance reduction. The difference of two consecutive RE linear combinations improves the asymptotic rate.

Definition 6.17 (RE estimator)**.** We define the *RE estimator* with coupling q as:

$$
\widehat{\mu}^{\mathrm{RE}\,q}_{v^{L,q}} := \sum_{\ell=1}^{L}\sum_{j=1}^{L}(v_j^{\ell,q}-v_j^{\ell-1,q})\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_j^{i,\ell}. \tag{6.22}
$$

◇

At first glance it looks like $\widehat{\mu}^{\mathrm{RE}\,q}_{v^{L,q}}$ uses all models on every level, however this is not true since some entries of the difference $v^{\ell,q}-v^{\ell-1,q}$ are zero. More precisely, the RE estimator with coupling q uses the following model groups

$$
S^\ell := \{\max\{\ell-q+1,1\},\max\{\ell-q+2,1\},\dots,\ell\} \quad \text{for } \ell\in\{1,\dots,L\} \tag{6.23}
$$

and thus couples at most q models. It is straightforward to verify that $\widehat{\mu}^{\mathrm{RE}\,q}_{v^{L,q}}$ is actually a linear unbiased estimator for $v^{L,q}$ since this estimator is constructed from a telescoping sum similar to MLMC. Furthermore, we write down the coefficients of the RE estimator in the form of this thesis

$$
\beta^\ell_{[j]} = v_j^{\ell,q}-v_j^{\ell-1,q} \quad \text{for all } j\in S^\ell, \quad \ell\in\{1,\dots,L\}. \tag{6.24}
$$

For all other model groups the coefficient is equal to zero. We further remark that for $q=2$ the RE and the MLMC estimator coincide.

Example 6.18 (RE 2 and MLMC are equal)**.** First, both the RE 2 and the MLMC estimator use the same model groups, which can be deduced by comparing (6.23) with (4.38). We write down the bias and coefficients of the RE estimator

$$\begin{aligned} v^{L,2} &= Dv^{L-1,2} = \cdots = D^{L-1}e_1 = e_L, \\ v^{\ell,2} - v^{\ell-1,2} &= e_\ell - e_{\ell-1}, \end{aligned}$$

where we define $e_0 := 0$. Therefore the coefficients as well as the bias of the RE 2 and the MLMC estimator are equal. In particular, the linear combination of the former satisfies

$$\sum_{j=1}^{L}(v_j^{\ell,q} - v_j^{\ell-1,q})\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} Z_j^{i,\ell} = \frac{1}{m_\ell}\sum_{i=1}^{m_\ell}(Z_\ell^{i,\ell} - Z_{\ell-1}^{i,\ell})$$

showing that both estimators are in fact equal. ◇

We verify that the bias and coefficients of the RE estimator lead to an improved bias and variance reduction rates if Assumption 6.10 or Assumption 6.11 are satisfied.

Lemma 6.19 (Properties of RE estimator coefficients)**.** Let Assumption 6.10 be satisfied. Then the RE vector has the bias rate $\gamma_{\text{Bias}} = \gamma_{q_{\text{mean}}}$

$$|(v^{L,q_{\text{mean}}})^T\mu - \mathbb{E}[Z]| \leq c2^{-\gamma_{q_{\text{mean}}}L}. \tag{6.25}$$

Now, let Assumption 6.11 be true. Then the difference of RE vectors has the variance reduction rate $\gamma_{\text{Var}} = 2\gamma_{q_{\text{path}}}$

$$(v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}})^T C(v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}) \leq c2^{-2\gamma_{q_{\text{path}}}\ell} \quad \text{for all } \ell \in \{1,\dots,L\}. \tag{6.26}$$

Furthermore, if the cost increase per level is geometrically bounded $w_\ell \leq c2^{\gamma_{\text{Cost}}\ell}$ for all $\ell \in \{1,\dots,L\}$, then so are the costs of the model groups

$$W^\ell \leq c2^{\gamma_{\text{Cost}}\ell} \quad \text{for all } \ell \in \{1,\dots,L\}. \tag{6.27}$$

Proof. The proof of (6.25) follows from (6.19) in Lemma 6.13 by subtracting the mean $\mathbb{E}[Z]$ from both sides

$$((v^{L,q_{\text{mean}}})^T\mu - \mathbb{E}[Z]) = \sum_{j=L+1}^{q_{\text{mean}}-1} c_j^L 2^{-\gamma_j\ell_0} + \mathcal{O}(2^{-\gamma_{q_{\text{mean}}}\ell_0}).$$

For $L \geq q_{\text{mean}} - 1$ the sum is zero. Since the first non–zero entry of $v^{L,q_{\text{mean}}}$ is the starting level $\ell_0 = L - q_{\text{mean}} + 1$ we conclude

$$|(v^{L,q_{\text{mean}}})^T\mu - \mathbb{E}[Z]| \leq c2^{-\gamma_{q_{\text{mean}}}(L-q_{\text{mean}}+1)} \leq c2^{-\gamma_{q_{\text{mean}}}L}.$$

The statement (6.25) also holds for $L < q_{\text{mean}} - 1$ if the constant c is large enough. We now prove (6.26). We use (6.20) in Lemma 6.14 to remove Z

$$\begin{aligned} &(v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}})^T C(v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}) \\ &= \mathbb{V}\left[\sum_{j=1}^{\ell} v_j^{\ell,q_{\text{path}}} Z_{j+\ell_0} - \sum_{j=1}^{\ell-1} v_j^{\ell-1,q_{\text{path}}} Z_{j+\ell_0}\right] \\ &= \mathbb{V}\left[\sum_{j=\ell+1}^{q_{\text{path}}-1} c_j^\ell 2^{-\gamma_j\ell_0} - \sum_{j=\ell}^{q_{\text{path}}-1} c_j^{\ell-1}2^{-\gamma_j\ell_0} + \mathcal{O}(2^{-\gamma_{q_{\text{path}}}\ell_0})\right]. \end{aligned}$$

For $\ell > q_{\text{path}} - 1$ the two sums disappear and using $\ell_0 = \ell - q_{\text{path}} + 1$ then shows (6.26). Once again, for $\ell \leq q_{\text{path}} - 1$ the constant has to be chosen sufficiently large. Finally, verifying (6.27) is straightforward due the geometric cost increase and the model groups defined in (6.23). □

We are now in the position to write down the complexity of the RE estimator.

Theorem 6.20 (Complexity of RE estimator)**.** Let both Assumption 6.10 and Assumption 6.11 be true with $q = q_{\text{mean}} = q_{\text{path}}$. Furthermore, assume a geometric cost increase of the models

$$w_\ell \leq c2^{\gamma_{\text{Cost}}\ell} \quad \text{for all } \ell \in \{1, \dots, L\}.$$

Then for all $\varepsilon \in (0, 1/e]$ there exists L and $m_1, \dots, m_L$ such that the RE estimator achieves $\text{MSE}[\widehat{\mu}^{\text{RE}\,q}_{v^{L,q}}] \leq \varepsilon^2$ with costs bounded by

$$\mathbb{W}\left[\widehat{\mu}^{\text{RE}\,q}_{v^{L,q}}\right] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_q} + c\begin{cases}\varepsilon^{-2}, & \text{if } 2\gamma_q > \gamma_{\text{Cost}},\\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } 2\gamma_q = \gamma_{\text{Cost}},\\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-2\gamma_q}{\gamma_q}}, & \text{if } 2\gamma_q < \gamma_{\text{Cost}}.\end{cases} \tag{6.28}$$

Proof. We have to verify the assumptions of Theorem 6.1. These however, follow either from the definition of the RE estimator in case of (6.5) and (6.6), or from Lemma 6.19 in case of (6.4), (6.7) and (6.8). The result (6.28) is then (6.9). □

For completeness we write down the corresponding upper cost bound for the SAOB.

Corollary 6.21 (RE complexity bound for the SAOB and SAOB κ)**.**
Let Assumption 6.10 and Assumption 6.11 be true with $q = q_{\text{mean}} = q_{\text{path}}$. Then for all $\varepsilon \in (0, 1/e]$ the SAOB and SAOB κ with $\kappa \geq q$ achieve a MSE of ε^2 with the cost bound in (6.28).

Weighted RE estimator. The bias and variance reduction capabilities of the RE estimator are coupled together in the sense that $\widehat{\mu}^{\text{RE}\,q}_{v^{L,q}}$ typically has equal rates $q_{\text{mean}} = q_{\text{path}}$. This is reasonable, however, we have already seen in Example 6.12 that necessarily $q_{\text{mean}} \geq q_{\text{path}}$ possibly with strict inequality. We combine this with $\gamma_{\text{Bias}} = \gamma_{q_{\text{mean}}}$ and $\gamma_{\text{Var}} = 2\gamma_{q_{\text{path}}}$ from Lemma 6.19 to obtain

$$2\gamma_{\text{Bias}} \geq \gamma_{\text{Var}}$$

with possibly strict inequality. Hence it is beneficial to apply RE more often for the mean than pathwise. Furthermore, we are interested in analysing the complexity of the $\widehat{\mu}^{\text{SAOB}}_L$ and $\widehat{\mu}^{\text{SAOB}\,q}_L$, which have a small bias $e_L = v^{L,2}$ but may have a large variance reduction rate comparable to $\widehat{\mu}^{\text{RE}\,q}_{v^{L,q}}$ for $q > 2$. We thus want to decouple the mean RE from the pathwise RE. To accommodate this change, we first observe that the difference of consecutive RE vectors form a basis of $\mathbb{R}^L$

$$\text{span}(v^{1,q_{\text{path}}} - v^{0,q_{\text{path}}}, \dots, v^{L,q_{\text{path}}} - v^{L-1,q_{\text{path}}}) = \mathbb{R}^L.$$

This follows from (6.18) observing that these vectors are linearly independent since

$$\begin{aligned} v^{\ell,q_{\text{path}}}_\ell &\neq 0 \quad \text{for all } \ell \in \{1, \dots, L\},\\ v^{\ell,q_{\text{path}}}_j &= 0 \quad \text{for all } j \in \{\ell+1, \dots, L\}, \quad \ell \in \{1, \dots, L\}. \end{aligned}$$

Hence, there exists weights $a := (a_1, \dots, a_L)^T \in \mathbb{R}^L$ such that for arbitrary $\alpha \in \mathbb{R}^L$

$$\alpha = \sum_{\ell=1}^{L} a_\ell (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}). \tag{6.29}$$

This allows us to define an RE estimator with arbitrary bias α.

Definition 6.22 (Weighted RE estimator)**.** For $\alpha \in \mathbb{R}^L$ we define the *weighted RE estimator* as

$$\widehat{\mu}_\alpha^{\text{RE}\,q} := \sum_{\ell=1}^{L} a_\ell \sum_{j=1}^{L} (v_j^{\ell,q} - v_j^{\ell-1,q}) \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_j^{i,\ell}, \tag{6.30}$$

where $a_1, \dots, a_L$ satisfy (6.29) with $q = q_{\text{path}}$. ◇

We verify that $\widehat{\mu}_\alpha^{\text{RE}\,q}$ is an unbiased estimator for $\alpha^T \mu$ using the definition of a in (6.29)

$$\mathbb{E}\big[\widehat{\mu}_\alpha^{\text{RE}\,q}\big] = \sum_{\ell=1}^{L} a_\ell \sum_{j=1}^{L} (v_j^{\ell,q} - v_j^{\ell-1,q}) \mu_j = \sum_{\ell=1}^{L} a_\ell (v^{\ell,q} - v^{\ell-1,q})^T \mu = \alpha^T \mu.$$

We verify that for $\alpha = v^{\ell,q}$ the weighted RE estimator (6.30) is equal to the standard RE estimator (6.22). We use the telescoping sum idea with $a_1 = \cdots = a_L = 1$ and $v^{0,q} = 0$ to conclude (6.29)

$$v^{\ell,q} = \sum_{\ell=1}^{L} (v^{\ell,q} - v^{\ell-1,q}) = \sum_{\ell=1}^{L} a_\ell (v^{\ell,q} - v^{\ell-1,q}).$$

The variance reduction properties of the weighted RE estimator compared to the standard RE estimator only differ in the additional weights $a_1, \dots, a_L$. We verify that these are bounded. In particular, we want to replace α with the RE vector $v^{L,q_{\text{mean}}}$ in (6.29) to potentially achieve a better bias rate.

Lemma 6.23 (Bounded weights)**.** For $\alpha = v^{L,q_{\text{mean}}}$ the weights $a_1, \dots, a_L$ satisfying (6.29) are bounded

$$|a_\ell| \le c \|v^{q_{\text{mean}}-1,q_{\text{mean}}}\| \le c \quad \text{for all } \ell \in \{1, \dots, L\}, \tag{6.31}$$

where the constant c is independent of the finest level L. Now let Assumption 6.11 be true. Then for all $\ell \in \{1, \dots, L\}$ the variance is bounded

$$(a_\ell (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}))^T C (a_\ell (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}})) \le c a_\ell^2 2^{-\gamma q_{\text{path}} \ell}. \tag{6.32}$$

Proof. The expression (6.32) follows from Lemma 6.19. We now show (6.31). Computing a in (6.29) means solving a linear system where the RE difference vectors

$$\Delta v^\ell := v^{1,q_{\text{path}}} - v^{0,q_{\text{path}}}$$

are columns of a matrix

$$\begin{aligned} &\big(\Delta v^1| \quad \dots \quad |\Delta v^L\big) a \\ &= \big(\Delta v^1| \quad \dots \quad |\Delta v^{q_{\text{path}}}| \quad D\Delta v^{q_{\text{path}}}| \quad D^2 \Delta v^{q_{\text{path}}}| \quad \dots \quad |D^{L-q_{\text{path}}} \Delta v^{q_{\text{path}}}\big) a = v^{L,q_{\text{mean}}}. \end{aligned}$$

Here we assume that L is sufficiently large. This matrix is upper triangular with non–zero diagonal and since the lower right part is independent of L, the values a_ℓ for all $\ell \in \{L - q_{\text{mean}} + 1, \dots, L\}$ are bounded independently of L

$$|a_\ell| \le c \|v^{L,q_{\text{mean}}}\| = c \|D^{L-q_{\text{mean}}+1} v^{q_{\text{mean}}-1,q_{\text{mean}}}\| = c \|v^{q_{\text{mean}}-1,q_{\text{mean}}}\|.$$

Now let $\ell \in \{q_{\text{path}}+1,\dots,L-q_{\text{mean}}\}$. We have that $v_\ell^{L,q_{\text{mean}}} = 0$ and since the first $L-q_{\text{path}}$ entries of $\Delta v^\ell = D^{\ell-q}v^q$ are zero due to the shift D, we obtain a_ℓ by a backward substitution with zero right–hand side

$$\Delta v_{q_{\text{path}}}^{q_{\text{path}}} a_\ell = -\sum_{j=1}^{q_{\text{path}}-1} \Delta v_{q_{\text{path}}-j}^{q_{\text{path}}} a_{\ell+j}.$$

Therefore the following recursive relationship is valid

$$\begin{pmatrix} a_\ell \\ \vdots \\ a_{\ell+q_{\text{path}}-1} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\Delta v_{q_{\text{path}}}^{q_{\text{path}}}} (\Delta v_{(q_{\text{path}}-1,\dots,1)}^{q_{\text{path}}})^T & 0 \\ I_{q_{\text{path}}-1,q_{\text{path}}-1} & 0 \end{pmatrix} \begin{pmatrix} a_{\ell+1} \\ \vdots \\ a_{\ell+q_{\text{path}}} \end{pmatrix} = Q \begin{pmatrix} a_{\ell+1} \\ \vdots \\ a_{\ell+q_{\text{path}}} \end{pmatrix}.$$

Here we use the notation that $\Delta v_{(q_{\text{path}}-1,\dots,1)}^{q_{\text{path}}} = (\Delta v_{q_{\text{path}}-1}^{q_{\text{path}}},\dots,\Delta v_1^{q_{\text{path}}})^T$ and suitably defined Q. We use this recursion to obtain an explicit term

$$a_\ell = e_1^T Q^{L-\ell-q_{\text{path}}+1} \begin{pmatrix} a_{L-q_{\text{path}}+1} \\ \vdots \\ a_L \end{pmatrix}.$$

We now show that $Q = U\Sigma U^{-1}$ is diagonalizable with diagonal matrix Σ with eigenvalues $\lambda_1,\dots,\lambda_L$ smaller or equal to one. Then $\|\Sigma^{L-\ell-q_{\text{path}}+1}\| \le 1$ and thus

$$\begin{aligned} |a_\ell| &\le \|e_1\| \|Q^{L-\ell-q_{\text{path}}+1}\| \left\| \begin{pmatrix} a_{L-q_{var}+1} \\ \vdots \\ a_L \end{pmatrix} \right\| \le c\|U\Sigma^{L-\ell-q_{\text{path}}+1}U^{-1}\| \|v^{q_{\text{mean}}-1,q_{\text{mean}}}\| \\ &\le c\|v^{q_{\text{mean}}-1,q_{\text{mean}}}\| \|U\| \|\Sigma^{L-\ell-q_{\text{path}}+1}\| \|U^{-1}\| \le c\|v^{q_{\text{mean}}-1,q_{\text{mean}}}\|. \end{aligned} \tag{6.33}$$

Clearly 0 is an eigenvalue of Q. We verify that the remaining $q_{\text{path}}-1$ distinct eigenpairs of Q are

$$(\lambda_j, x^j) := (2^{-\gamma_j}, (2^{-\gamma_j(q_{\text{path}}-n+1)})_{n=1}^{q_{\text{path}}}) \qquad \text{for all } j \in \{1,\dots,q_{\text{path}}-1\}.$$

We compute the application of Q

$$\begin{aligned} Qx^j &= \begin{pmatrix} -\frac{1}{\Delta v_{q_{\text{path}}}^{q_{\text{path}}}} (\Delta v_{(q_{\text{path}}-1,\dots,1)}^{q_{\text{path}}})^T & 0 \\ I_{q_{\text{path}}-1,q_{\text{path}}-1} & 0 \end{pmatrix} \begin{pmatrix} 2^{-\gamma_j q_{\text{path}}} \\ \vdots \\ 2^{-\gamma_j} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{\Delta v_{q_{\text{path}}}^{q_{\text{path}}}} (\Delta v_{(q_{\text{path}}-1,\dots,1)}^{q_{\text{path}}})^T x^j_{(1,\dots,q_{\text{path}}-1)} \\ 2^{-\gamma_j q_{\text{path}}} \\ 2^{-\gamma_j(q_{\text{path}}-1)} \\ \vdots \\ 2^{-\gamma_j 2} \end{pmatrix}. \end{aligned}$$

Clearly, the equation $Qx^j = 2^{-\gamma_j}x^j$ is true if the first row satisfies

$$-(\Delta v_{(q_{\text{path}}-1,\dots,1)}^{q_{\text{path}}})^T x^j_{(1,\dots,q_{\text{path}}-1)} = \Delta v_{q_{\text{path}}}^{q_{\text{path}}} 2^{-\gamma_j} 2^{-\gamma_j q_{\text{path}}}.$$

We divide both sides by $2^{-2\gamma_j}$, rearrange the terms and thus we have to show that $2^{-\gamma_j}$ is a root of the polynomial

$$\sum_{n=1}^{q_{\text{path}}} \Delta v_n^{q_{\text{path}}} 2^{-\gamma_j(n-1)} = 0. \tag{6.34}$$

We compute the RE difference $\Delta v^{q_{\text{path}}} = D v^{q_{\text{path}}-1,q_{\text{path}}} - v^{q_{\text{path}}-1,q_{\text{path}}}$ with the help of the explicit form (6.21), where a calculation and rewriting as multiplication shows

$$\begin{aligned}
\sum_{n=1}^{q_{\text{path}}} (D v^{q_{\text{path}}-1,q_{\text{path}}})_n 2^{-\gamma_j(n-1)} &= \sum_{n=2}^{q_{\text{path}}} v_{n-1}^{q_{\text{path}}-1,q_{\text{path}}} 2^{-\gamma_j(n-1)} \\
&= \frac{1}{\prod_{n=2}^{q_{\text{path}}-1}(2^{\gamma_n}-1)} \prod_{n=2}^{q_{\text{path}}-1} (1 - 2^{\gamma_n} 2^{-\gamma_j}) \\
\sum_{n=1}^{q_{\text{path}}} v_n^{q_{\text{path}}-1,q_{\text{path}}} 2^{-\gamma_j(n-1)} &= 2^{-\gamma_j} \sum_{n=1}^{q_{\text{path}}-1} v^{q_{\text{path}}-1,q_{\text{path}}{}_n} 2^{-\gamma_j n} \\
&= \frac{2^{-\gamma_j}}{\prod_{n=2}^{q_{\text{path}}-1}(2^{\gamma_n}-1)} \prod_{n=2}^{q_{\text{path}}-1} (1 - 2^{\gamma_n} 2^{-\gamma_j}).
\end{aligned}$$

The condition (6.34) is thus equivalent to

$$\prod_{n=2}^{q_{\text{path}}-1} (1 - 2^{\gamma_n} 2^{-\gamma_j}) = 2^{-\gamma_j} \prod_{n=2}^{q_{\text{path}}-1} (1 - 2^{\gamma_n} 2^{-\gamma_j}).$$

Equality now trivially holds for $j = 2, \dots, q_{\text{path}} - 1$ since then the j–th factor is equal to zero. For $j = 1$ we have $\gamma_1 = 0$ by definition and thus $2^{-\gamma_1} = 1$ showing equality. We conclude (6.33) and summarize the result so far

$$|a_\ell| \le c \| v^{q_{\text{mean}}-1,q_{\text{mean}}} \| \qquad \text{for all } \ell \in \{q_{\text{path}}+1, \dots, L\},$$

where the constant c is independent of L. Finally, the weights $a_1, \dots, a_{q_{\text{path}}}$ depend linearly on $a_{q_{\text{path}}+1}, \dots, a_{2q_{\text{path}}}$ and the matrix describing this dependence is independent of L. We conclude (6.31) and thus the lemma. □

Let us formulate a corollary for the boundedness of the weights for arbitrary $\alpha \in \mathbb{R}^L$. This allows us to use an arbitrary bias vector, even if the bias vector is not an RE vector.

Corollary 6.24 (Bounded weights)**.** For $\alpha \in \mathbb{R}^L$ and $a_1, \dots, a_L$ satisfying (6.29) the bound holds

$$|a_\ell| \le c \|\alpha\|_{\ell^1} \quad \text{for all } \ell \in \{1, \dots, L\},$$

where the constant c is independent of L. In particular, if a sequence of bias vectors $(\alpha^L)_{L=1}^\infty$ is constructed from a down shift $\alpha^L := D^{L-q}\alpha^q$ for some α^q and some fixed q, i.e. the RE vector $v^{L,q}$, then

$$|a_\ell| \le c \quad \text{for all } \ell \in \{1, \dots, L\} \text{ and all } L \in \mathbb{N}.$$

Proof. We express the bias in the basis of unit vectors, which are RE vectors

$$\alpha = \sum_{j=1}^{L} b_j e_j = \sum_{j=1}^{L} b_j v^{j,2}. \tag{6.35}$$

The idea of the proof is now a component–by–component application of Lemma 6.23. Define weights a^j that satisfy (6.29)

$$v^{j,2} = \sum_{\ell=1}^{L} a_\ell^j (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}) \quad \text{for all } j \in \{1,\dots,L\}.$$

Setting $a_\ell^j = 0$ since $v_\ell^{j,2} = 0$ for $\ell > j$ shows

$$v^{j,2} = \sum_{\ell=1}^{j} a_\ell^j (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}) \quad \text{for all } j \in \{1,\dots,L\}. \tag{6.36}$$

This system has a solution according to (6.31) in Lemma 6.23 with bounded components

$$|a_\ell^j| \leq c\|v^{j,2}\| = c\|e_j\| = c. \tag{6.37}$$

We use (6.35) together with (6.36)

$$\alpha = \sum_{j=1}^{L} b_j v^{j,2} = \sum_{j=1}^{L} b_j \sum_{\ell=1}^{L} a_\ell^j (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}) = \sum_{\ell=1}^{L} \left(\sum_{j=1}^{L} b_j a_\ell^j \right) (v^{\ell,q_{\text{path}}} - v^{\ell-1,q_{\text{path}}}),$$

hence a comparison of coefficients shows that the weights satisfy $a_\ell = \sum_{j=1}^{L} b_j a_\ell^j$. We combine the bound (6.37) with (6.35) to conclude the corollary

$$|a_\ell| = \left| \sum_{j=1}^{L} b_j a_\ell^j \right| \leq \sum_{j=1}^{L} |b_j||a_\ell^j| \leq c \sum_{j=1}^{L} |b_j| = c\|\alpha\|_{\ell^1}. \qquad \square$$

Lemma 6.23 allows us to write down the complexity of the weighted RE estimator.

Theorem 6.25 (Complexity of weighted RE estimator)**.** Let Assumption 6.10 and Assumption 6.11 be true. Furthermore, assume a geometric cost increase of the models

$$w_\ell \leq c 2^{\gamma_{\text{Cost}} \ell} \quad \text{for all } \ell \in \{1,\dots,L\}.$$

Then for all $\varepsilon \in (0, 1/e]$ there exists L and $m_1,\dots,m_L$ such that $\text{MSE}[\widehat{\mu}_{v^{L},q_{\text{mean}}}^{\text{RE}\,q_{\text{path}}}] \leq \varepsilon^2$ with costs bounded by

$$\mathbb{W}\left[\widehat{\mu}_{v^{L},q_{\text{mean}}}^{\text{RE}\,q_{\text{path}}}\right] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{q_{\text{mean}}}} + c \begin{cases} \varepsilon^{-2}, & \text{if } 2\gamma_{q_{\text{path}}} > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } 2\gamma_{q_{\text{path}}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-2\gamma_{q_{\text{path}}}}{\gamma_{q_{\text{mean}}}}}, & \text{if } 2\gamma_{q_{\text{path}}} < \gamma_{\text{Cost}}. \end{cases} \tag{6.38}$$

Proof. The proof is to combine Lemma 6.19 and Lemma 6.23 to verify Theorem 6.1. $\square$

We write down the corollary for the SAOBs.

Corollary 6.26 (Weighted RE complexity bound for SAOB and SAOB κ)**.**
Let Assumption 6.10 and Assumption 6.11 be true. Then for all $\varepsilon \in (0, 1/e]$ the estimators $\widehat{\mu}^{\mathrm{SAOB}}_{v^{L,q_{\mathrm{mean}}}}$ and $\widehat{\mu}^{\mathrm{SAOB}\,\kappa}_{v^{L,q_{\mathrm{mean}}}}$ with $\kappa \geq q_{\mathrm{path}}$ achieve a MSE of ε^2 with the cost bound in (6.38).

Analytic complexity. Both Assumption 6.10 and Assumption 6.11 have a finite index q_{mean} or q_{path} respectively, where the series expansion ends with a remainder. It is now straightforward to verify that a RE estimator with sufficiently high order achieves optimal complexity. For simplicity, we only state the result with the stronger pathwise assumption.

Theorem 6.27 (Analytic complexity of RE estimator)**.** Let $(\gamma_n)_{n=1}^{\infty}$ be a sequence with $0 = \gamma_1 < \gamma_2 < \dots$ and $\gamma_n \to +\infty$ for $n \to +\infty$ such that $\mathbb{P}$–almost surely

$$Z_\ell(\omega) = Z(\omega) + \sum_{j=2}^{\infty} c_j(\omega) 2^{-\gamma_j \ell},$$

where $c_2, c_3, \dots$ have finite second moments uniformly bounded by some constant c. Furthermore, assume that the models satisfy the geometric cost bound

$$w_\ell \leq c 2^{\gamma_{\mathrm{Cost}} \ell} \qquad \text{for all } \ell \in \mathbb{N}.$$

Then for all $\varepsilon \in (0, 1/e]$ there exists q, L and $m_1, \dots, m_L$ such that $\mathrm{MSE}[\widehat{\mu}^{\mathrm{RE}\,q}_{v^{L,q}}] \leq \varepsilon^2$ with costs bounded by

$$\mathbb{W}\left[\widehat{\mu}^{\mathrm{RE}\,q}_{v^{L,q}}\right] \leq c\varepsilon^{-2}.$$

Proof. We apply Theorem 6.20 with sufficiently large q. Since $\gamma_n \to +\infty$ we choose q large enough such that both the rounding costs and the variance costs are smaller or equal to ε^{-2}. This is achieved if we ensure that

$$\begin{aligned} -\gamma_{\mathrm{Cost}}/\gamma_q &\geq -2, \\ 2\gamma_q &> \gamma_{\mathrm{Cost}}, \end{aligned}$$

which is satisfied if the second condition is true $2\gamma_q > \gamma_{\mathrm{Cost}}$. □

The SAOB is then also asymptotically optimal with costs of order ε^{-2}.

Corollary 6.28 (Analytic RE complexity bound for SAOB and SAOB κ)**.** Let the assumptions of Theorem 6.27 be true. Then for all $\varepsilon \in (0, 1/e]$ there exist L, q such that the estimators $\widehat{\mu}^{\mathrm{SAOB}}_{v^{L,q}}$ and $\widehat{\mu}^{\mathrm{SAOB}\,\kappa}_{v^{L,q}}$ with $\kappa \geq q$ achieve a MSE of ε^2 with costs bounded by $c\varepsilon^{-2}$.

6.3 Lower bounds on the complexity

The goal of this section is to derive lower bounds on the complexity of the SAOB. First of all, if the complexity of the RE estimator has already the optimal costs of order ε^{-2} of any sampling based estimator, then this is clearly also a lower bound of SAOB. The other interesting case is when the costs on the finest level dominate. Unsurprisingly, if the costs on the fine grid dominates, then we may assume low fidelity samples are for free without increasing the asymptotic complexity. The basic idea in this section is thus to use the lower variance bound $\mathbb{V}^{\min}$ from Definition 3.19 to obtain a lower bound on the complexity for the SAOB.

Lemma 6.29 (Lower bound for costs)**.** Let the bias $\alpha_L \neq 0$ and the covariance matrix C be positive definite. Then the cost of every linear unbiased estimator to reach a variance of ε^2 is lower bounded by

$$\mathbb{W}[\widehat{\mu}_\alpha] \geq c\varepsilon^{-2}\alpha_L^2 \mathbb{W}[Z_L]\mathbb{V}^{\min}. \tag{6.39}$$

Proof. For fixed coefficients β^k the cost to reach a variance of ε^2 is (6.3)

$$\mathbb{W}[\widehat{\mu}_\alpha] = c\varepsilon^{-2}\left(\sum_{k=1}^{K}\left((\beta^k)^T C^k \beta^k W^k\right)^{1/2}\right)^2.$$

This expression does not include ceiling the number of samples, which we neglect since we look for a lower bound. We minimize the inner expression if we introduce more degrees of freedom by replacing $\beta^k \in \mathbb{R}^{|S^k|}$ with $\beta^k \in \mathbb{R}^L$, $C^k \in \mathbb{R}^{|S^k|\times|S^k|}$ with $C \in \mathbb{R}^{L\times L}$ and setting $w_\ell = 0$ for $\ell = 1, \dots, L-1$. We further assume now $\beta_L^k = 0$ if $L \notin S^k$ and we further relax the bias constraint to only be enforced on the finest level

$$\left(\sum_{k=1}^{K}\left((\beta^k)^T C^k \beta^k W^k\right)^{1/2}\right)^2 \geq w_L\left(\sum_{\{k\,|\,L\in S^k\}}\left((\beta^k)^T C \beta^k\right)^{1/2}\right)^2$$
$$\alpha_L = \sum_{\{k\,|\,L\in S^k\}} \beta_L^k.$$

We view this now as minimization problem, remark that every summand contains the same value and appears equally in the equality constraint and hence we w.l.o.g. assume $(\beta^k)^T C \beta^k \neq 0$. We minimize over the convex inner expression to obtain the KKT conditions for some Lagrange-multiplier $\lambda \in \mathbb{R}$

$$\frac{C\beta^k}{\left((\beta^k)^T C\beta^k\right)^{1/2}} = \lambda e_L \qquad \text{for all } k \text{ with } L \in S^k,$$
$$\alpha_L = \sum_{\{k\,|\,L\in S^k\}} \beta_L^k.$$

From the first condition we conclude

$$\frac{\beta^k}{\left((\beta^k)^T C\beta^k\right)^{1/2}} = \frac{\beta^j}{\left((\beta^j)^T C\beta^j\right)^{1/2}} \qquad \text{for all } k, j \text{ with } L \in S^k \cap S^j.$$

We are now able to choose these $\beta^k = \beta \neq 0$ independently of k such that $(\beta^k)^T C\beta^k = \beta^T C\beta$. The bias constraint can also be satisfied if

$$\alpha_L = \sum_{\{k\,|\,L\in S^k\}} \beta_L^k = |\{k\,|\,L \in S^k\}|\beta_L.$$

Therefore, we are allowed to choose $\beta_1, \dots, \beta_{L-1}$ freely and only β_L is fixed. We use the last equation to conclude

$$\left(\sum_{\{k\,|\,L\in S^k\}}\left((\beta^k)^T C\beta^k\right)^{1/2}\right)^2 \geq \left(|\{k\,|\,L\in S^k\}|\left(\beta^T C\beta\right)^{1/2}\right)^2 = |\{k\,|\,L\in S^k\}|^2\beta^T C\beta$$
$$= \begin{pmatrix}\widetilde{\beta}\\ \alpha_L\end{pmatrix}^T C \begin{pmatrix}\widetilde{\beta}\\ \alpha_L\end{pmatrix} = \alpha_L^2 \begin{pmatrix}\overline{\beta}\\ 1\end{pmatrix}^T C \begin{pmatrix}\overline{\beta}\\ 1\end{pmatrix}.$$

Since we minimize over $\overline{\beta}$ Lemma 4.17 shows that the last expression is equal to the lower variance bound scaled by α_L^2 and thus

$$\mathbb{W}[\widehat{\mu}_\alpha] \geq c\varepsilon^{-2} w_L \alpha_L^2 \mathbb{V}^{\min}. \qquad \square$$

It is often possible to derive asymptotic bounds for α_L and $\mathbb{W}[Z_L]$, e.g. using RE vectors to achieve a certain bias rate and assuming a geometric cost increase due to the discretization of Z. However, an asymptotic expression and lower bound for $\mathbb{V}^{\min}$ is not obvious. We first derive an upper bound for $\mathbb{V}^{\min}$.

Lemma 6.30 (Asymptotic upper bound for $\mathbb{V}^{\min}$)**.** Let Assumption 6.11 be true. Then

$$\mathbb{V}^{\min} \leq c2^{-2L\gamma_{q_{\text{path}}}}. \tag{6.40}$$

Proof. We use the definition of $\mathbb{V}^{\min}$ as minimizer (4.30) and insert the (suboptimal) difference of RE vectors, which we scale in the last component using

$$v_L^{L,q_{\text{path}}} - v_L^{L-1,q_{\text{path}}} = v_L^{L,q_{\text{path}}}.$$

We then use (6.26) to asymptotically bound the variance

$$\min_{\beta\in\mathbb{R}^{L-1}} \mathbb{V}\left[Z_L + \sum_{\ell=1}^{L-1} \beta_\ell Z_\ell\right] \leq \frac{(v^{L,q_{\text{path}}} - v^{L-1,q_{\text{path}}})^T C (v^{L,q_{\text{path}}} - v^{L-1,q_{\text{path}}})}{(v_L^{L,q_{\text{path}}})^2} \leq c2^{-2L\gamma_{q_{\text{path}}}}.$$

$\square$

We give an example where the rate (6.40) is actually sharp.

Example 6.31 (Expansion with independent noise for $\mathbb{V}^{\min}$)**.** Define the QoI as follows

$$Z_\ell(\omega) := Z(\omega) + \sum_{j=2}^{q_{\text{path}}-1} c_j(\omega) 2^{-\gamma_j \ell} + \xi_\ell(\omega) 2^{-\gamma_{q_{\text{path}}} \ell}.$$

We assume that $Z, c_2, \dots, c_{q_{\text{path}}-1}, \xi_1, \dots, \xi_\ell \sim N(0,1)$ are independent random variables and hence Assumption 6.11 is satisfied. Then every linear combination of models satisfies

$$\mathbb{V}\left[\sum_{\ell=1}^{L} \beta_\ell Z_\ell\right] = \left(\sum_{\ell=1}^{L} \beta_\ell\right)^2 \mathbb{V}[Z] + \sum_{j=2}^{q_{\text{path}}-1} \left(\sum_{\ell=1}^{L} \beta_\ell 2^{-\gamma_j \ell}\right)^2 \mathbb{V}[c_j] + \sum_{\ell=1}^{L} \beta_\ell^2 2^{-2\gamma_{q_{\text{path}}} \ell} \mathbb{V}[\xi_\ell].$$

The first two summands are zero for the difference of RE coefficients that use RE sufficiently often. We know $\beta_L = 1$ and thus we lower bound the last term

$$\mathbb{V}\left[\sum_{\ell=1}^{L} \beta_\ell Z_\ell\right] \geq \beta_L^2 2^{-2L\gamma_{q_{\text{path}}}} = 2^{-2L\gamma_{q_{\text{path}}}}.$$

We conclude $\mathbb{V}^{\min} \geq c2^{-2L\gamma_{q_{\text{path}}}}$ and the sharpness of (6.40). $\diamond$

Now that we have shown the underlying statement for the variance we adapt this result for the mean.

Example 6.32 (Expansion with independent noise for bias)**.** Define the QoI as follows

$$Z_\ell(\omega) := Z(\omega) + \sum_{j=2}^{q_{\text{path}}-1} c_j(\omega) 2^{-\gamma_j \ell} + \xi_\ell(\omega) 2^{-\gamma_{q_{\text{path}}} \ell}.$$

We assume that $Z, c_2, \dots, c_{q_{\text{path}}-1}, \xi_1, \dots, \xi_\ell \sim N(c, 1)$ are independent random variables such that each one has an unknown mean. We then take the mean to obtain

$$\mathbb{E}[Z_\ell] = \mathbb{E}[Z] + \sum_{j=2}^{q_{\text{path}}-1} \mathbb{E}[c_j] 2^{-\gamma_j \ell} + \mathbb{E}[\xi_\ell] 2^{-\gamma_{q_{\text{path}}} \ell},$$

which satisfies Assumption 6.10 with $q_{\text{mean}} = q_{\text{path}}$. We linearly combine them to obtain

$$\sum_{\ell=1}^{L} \alpha_\ell \mathbb{E}[Z_\ell] - \sum_{\ell=1}^{L} \alpha_\ell \mathbb{E}[Z] = \sum_{j=2}^{q_{\text{mean}}-1} \mathbb{E}[c_j] \sum_{\ell=1}^{L} \alpha_\ell 2^{-\gamma_j \ell} + \sum_{\ell=1}^{L} \alpha_\ell \mathbb{E}[\xi_\ell] 2^{-\gamma_{q_{\text{mean}}} \ell}.$$

Under the assumption $\sum_{\ell=1}^{L} \alpha_\ell = 1$ we obtain a standard bias expression and using RE coefficients ensures that the first sum disappears. We have

$$\sum_{\ell=1}^{L} \alpha_\ell \mathbb{E}[Z_\ell] - \mathbb{E}[Z] = \sum_{\ell=1}^{L} \alpha_\ell \mathbb{E}[\xi_\ell] 2^{-\gamma_{q_{\text{mean}}} \ell}.$$

This result holds even if we do not know mean values of $\mathbb{E}[c_j]$. We however, do not know $\mathbb{E}[\xi_\ell]$, which prevents us from removing the last sum. In the worst case, we may have $\mathbb{E}[\xi_\ell] = \operatorname{sign}(\alpha_\ell)$, we use $\sum_{\ell=1}^{L} \alpha_\ell = 1$ and we simplify by assuming known means $\mathbb{E}[c_j] = 0$ to obtain the lower bound

$$\left| \sum_{\ell=1}^{L} \alpha_\ell \mathbb{E}[Z_\ell] - \mathbb{E}[Z] \right| = \left| \sum_{\ell=1}^{L} |\alpha_\ell| 2^{-\gamma_{q_{\text{mean}}} \ell} \right| \geq 2^{-\gamma_{q_{\text{mean}}} L} \sum_{\ell=1}^{L} |\alpha_\ell| \geq 2^{-L\gamma_{q_{\text{mean}}}}.$$

This example shows that in general, the rate for the bias cannot be better than $2^{-L\gamma_{q_{\text{mean}}}}$.

◇

We now state a main result concerning a lower bound of the asymptotic complexity. This result relies on the crucial assumption that the bias rate is lower bounded in terms of Assumption 6.10 and the lower variance bound $\mathbb{V}^{\min}$ in terms of Assumption 6.11. The two previous examples showed that in general we cannot improve this result.

Theorem 6.33 (Lower complexity bound for SAOB)**.** Let Assumption 6.10 and 6.11 be true and C be positive definite. Furthermore, assume that the bias bound, the lower variance bound and the lower cost bound

$$|(v^{L,q_{\text{mean}}})^T \mu - \mathbb{E}[Z]| \geq c 2^{-L\gamma_{q_{\text{mean}}}} \qquad \text{for all } L \in \mathbb{N}, \tag{6.41}$$

$$\mathbb{V}^{\min} \geq c 2^{-2L\gamma_{q_{\text{path}}}} \qquad \text{for all } L \in \mathbb{N}, \tag{6.42}$$

$$w_L \geq c 2^{L\gamma_{\text{Cost}}} \qquad \text{for all } L \in \mathbb{N}. \tag{6.43}$$

Then for all $\varepsilon \in (0, 1/e]$ the cost to achieve $\operatorname{MSE}[\widehat{\mu}^{\text{SAOB}}_{v^{L,q_{\text{mean}}}}] \leq \varepsilon^2$ is lower bounded by

$$\mathbb{W}\left[\widehat{\mu}^{\text{SAOB}}_{v^{L,q_{\text{mean}}}}\right] \geq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{q_{\text{mean}}}} + c \begin{cases} \varepsilon^{-2}, & \text{if } 2\gamma_{q_{\text{path}}} \geq \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-2\gamma_{q_{\text{path}}}}{\gamma_{q_{\text{mean}}}}}, & \text{if } 2\gamma_{q_{\text{path}}} < \gamma_{\text{Cost}}. \end{cases} \tag{6.44}$$

This bound is sharp for $2\gamma_{q_{\text{path}}} > \gamma_{\text{Cost}}$ and $2\gamma_{q_{\text{path}}} < \gamma_{\text{Cost}}$.

Proof. The first term in the bound (6.44) is simply the cost for a single high fidelity evaluation and thus cannot be lower bounded any further. This evaluation is required, since $v_L^{L,q_{\text{mean}}} \neq 0$. For $2\gamma_{q_{\text{path}}} \geq \gamma_{\text{Cost}}$ the bound of ε^{-2} is trivial. For the case $2\gamma_{q_{\text{path}}} < \gamma_{\text{Cost}}$ the bias-variance decomposition shows

$$\mathbb{E}\big[(\widehat{\mu}^{\text{SAOB}}_{v^{L,q_{\text{mean}}}} - \mathbb{E}[Z])^2\big] = ((v^{L,q_{\text{mean}}})^T\mu - \mathbb{E}[Z])^2 + \mathbb{V}\big[\widehat{\mu}^{\text{SAOB}}_{v^{L,q_{\text{mean}}}}\big].$$

To achieve a bias of order ε^2 we have to choose $L \geq -\log_2(\varepsilon)/\gamma_{q_{\text{mean}}} + c$ using (6.41). Since we want a tight lower bound, we choose $L = -\log_2(\varepsilon)/\gamma_{q_{\text{mean}}}$. We now combine this with (6.39), the bound (6.42) and (6.43)

$$\mathbb{W}[\widehat{\mu}_\alpha] \geq c\varepsilon^{-2}\alpha_L^2\mathbb{W}[Z_L]\mathbb{V}^{\min} \geq c\varepsilon^{-2}2^{L(\gamma_{\text{Cost}}-2\gamma_{q_{\text{path}}})} = c\varepsilon^{-2}\varepsilon^{-(\gamma_{\text{Cost}}-2\gamma_{q_{\text{path}}})/\gamma_{q_{\text{mean}}}}.$$

This shows the claim (6.44). The sharpness of the bound follows from Corollary 6.26. □

The previous theorem shows that there is a strict lower limit on the asymptotic complexity of any linear unbiased estimator. In particular, the smoothness of $Z_1, Z_2, \ldots$ w.r.t. the mean and pathwise can be used to offset the increased cost w.r.t. the level. However, this only works up to a certain degree and thus the optimal asymptotic costs of $c\varepsilon^{-2}$ cannot be achieved in general for linear unbiased estimators.

Remark 6.34 (Equal costs on every level)**.** The previous theorem does not include the case of a sharp bound for equal costs on every level $2\gamma_{q_{\text{path}}} = \gamma_{\text{Cost}}$, in which the proof idea of Lemma 6.29 falls apart. We cannot reduce

$$\left(\sum_{k=1}^{K}\big((\beta^k)^T C^k \beta^k W^k\big)^{1/2}\right)^2$$

to a single model group and we also cannot pull out the costs for only the high fidelity model. Intuitively, this is not possible since each non–zero term should asymptotically be of the same size. We thus expect

$$\left(\sum_{k=1}^{K}\big((\beta^k)^T C^k \beta^k W^k\big)^{1/2}\right)^2 \geq c(L(\beta^L)^T C^L \beta^L W^L)^{1/2})^2 = cL^2(\beta^L)^T C^L \beta^L W^L, \quad (6.45)$$

where S^L is now the only model group that contains the high fidelity model. It is however, not straightforward to verify this inequality. However, if (6.45) holds, then it is straightforward to provide a sharp bound for $2\gamma_{q_{\text{path}}} = \gamma_{\text{Cost}}$ following the ideas outlined in this section. The additional L^2 compared to (6.39) adds the term $\log(\varepsilon)^2$. In any case, it is straightforward to verify that under the assumption of Theorem 6.33 for $2\gamma_{q_{\text{path}}} = \gamma_{\text{Cost}}$ and using Corollary 6.26

$$c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{q_{\text{mean}}}} + c\varepsilon^{-2} \leq \mathbb{W}\big[\widehat{\mu}^{\text{SAOB}}_{v^{L,q_{\text{mean}}}}\big] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{q_{\text{mean}}}} + c\varepsilon^{-2}\log(\varepsilon)^2.$$

Therefore, the bound (6.44) is tight up to logarithmic factors. ◇

Remark 6.35 (C positive definite in Theorem 6.33)**.** Let us comment on the assumption that C in Theorem 6.33 is positive definite, which is not necessary. If C has a zero eigenvalue and corresponding eigenvector v with $v_L \neq 0$ then

$$\mathbb{V}^{\min} = \min_\beta \begin{pmatrix}\beta\\1\end{pmatrix}^T C \begin{pmatrix}\beta\\1\end{pmatrix} \leq \frac{1}{v_L^2} v^T C v = 0,$$

which contradicts (6.42). Therefore, C may only have zero eigenvalues with eigenvectors v where $v_L = 0$. This means that linear combinations of low fidelity models may be linearly dependent (ignoring the mean). However, since we assumed $w_\ell = 0$ for $\ell \in \{1, \dots, L-1\}$ we cannot use this linear dependence to further reduce the variance. Thus, we may always remove enough models of $Z_1, \dots, Z_{L-1}$ such that C is positive definite without increasing the variance. ◇

6.4 Numerical experiments with explicit expansions

Convergence of SAOB to RE. We demonstrate that for a specific example the SAOB converges to the RE estimator. We define the QoI as

$$Z_\ell(\omega) := Z(\omega) + \sum_{j=2}^{4} c_j 2^{-(j-1)(\ell+\ell_0)} + 0.1\xi_\ell 2^{-3(\ell+\ell_0)} \tag{6.46}$$

and assume that $(Z, c_2, c_3, c_4)^T \sim N(0, Q)$ with covariance matrix $Q_{ij} = \exp(-|i-j|)$, which is obtained from the exponential covariance function. We further assume that $\xi_\ell \sim N(0,1)$ i.i.d. and that the ξ_ℓ are independent of Z, c_2, c_3, c_4. The variable ℓ_0 describes the initial accuracy of the models. We compute the covariance matrix C analytically and fix artificial costs for every model $\ell \in \{1, \dots, L\}$ with $w_\ell = 2^{2\ell-2}$. We further choose the computational budget $\mathbb{W}^{\text{budget}}$ to be sufficiently large. This model fits Assumption 6.11 with $q_{\text{path}} = 4$, $\gamma_2 = 1$, $\gamma_3 = 2$ and $\gamma_4 = 3$. We compare the SAOB κ with the weighted RE κ estimator with $\kappa = 2, 3, 4$ with a bias of $\alpha := e_L$. We define the difference in the coefficients as follows

$$d_\kappa(\ell_0) := \left(\sum_{k=1}^{K} \|\beta^{k,\text{SAOB}\,\kappa}(\ell_0) - \beta^{k,\text{RE}\,\kappa}\|^2 \right)^{1/2}. \tag{6.47}$$

The SAOB chooses the same model groups as the RE estimators for example (6.46). We further compare the relative difference of the variances

$$v_\kappa(\ell_0) := \frac{\mathbb{V}\left[\widehat{\mu}_L^{\text{RE}\,\kappa}(\ell_0)\right] - \mathbb{V}\left[\widehat{\mu}_L^{\text{SAOB}\,\kappa}(\ell_0)\right]}{\mathbb{V}[\widehat{\mu}_L^{\text{SAOB}\,\kappa}(\ell_0)]}. \tag{6.48}$$

The variance of both estimators clearly depends on the covariance matrix $C = C(\ell_0)$ which itself depends on ℓ_0. However, the coefficients of the RE estimators $\beta^{k,\text{RE}\,\kappa}$ do not depend on ℓ_0. We plot the computed values d_κ and v_κ in Figure 6.1 for different values of ℓ_0. The respective coefficients of the SAOB 2 and the MLMC (RE 2) estimator are given in Figure 6.2 and for the SAOB 3 and RE 3 estimator in Figure 6.3. We conclude that both the coefficients as well as the relative variance difference converges to zero. This indicates that asymptotically, there is no difference between the SAOB κ and RE κ estimator. We further conclude that the difference between the SAOB κ and the RE κ estimator is larger if κ is larger, both in terms of d_κ and v_κ. This is not surprising, since the SAOB is able to optimize over more parameters achieving a larger variance reduction.

Remark 6.36 (Proof of convergence)**.** A rigorous proof of convergence for the models (6.46) is not straightforward as it might seem. An idea is to first verify that both estimators use the same model groups, i.e. for sufficiently large ℓ_0 the SAOB κ uses the

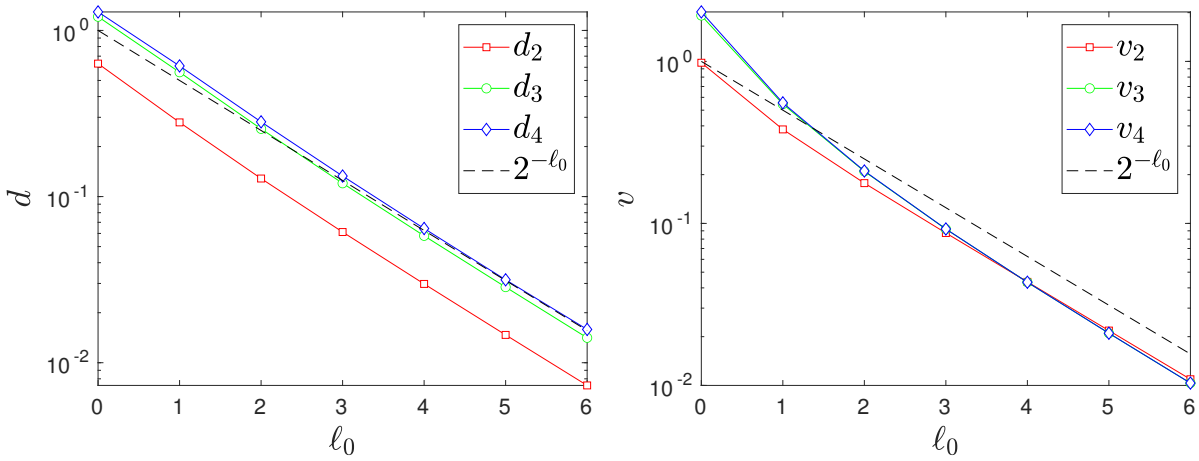

Figure 6.1: The left image shows the convergence of the coefficients of the SAOB κ to the coefficients of the weighted RE κ estimator w.r.t. (6.47). The right image shows the convergence of the relative variance difference (6.48).

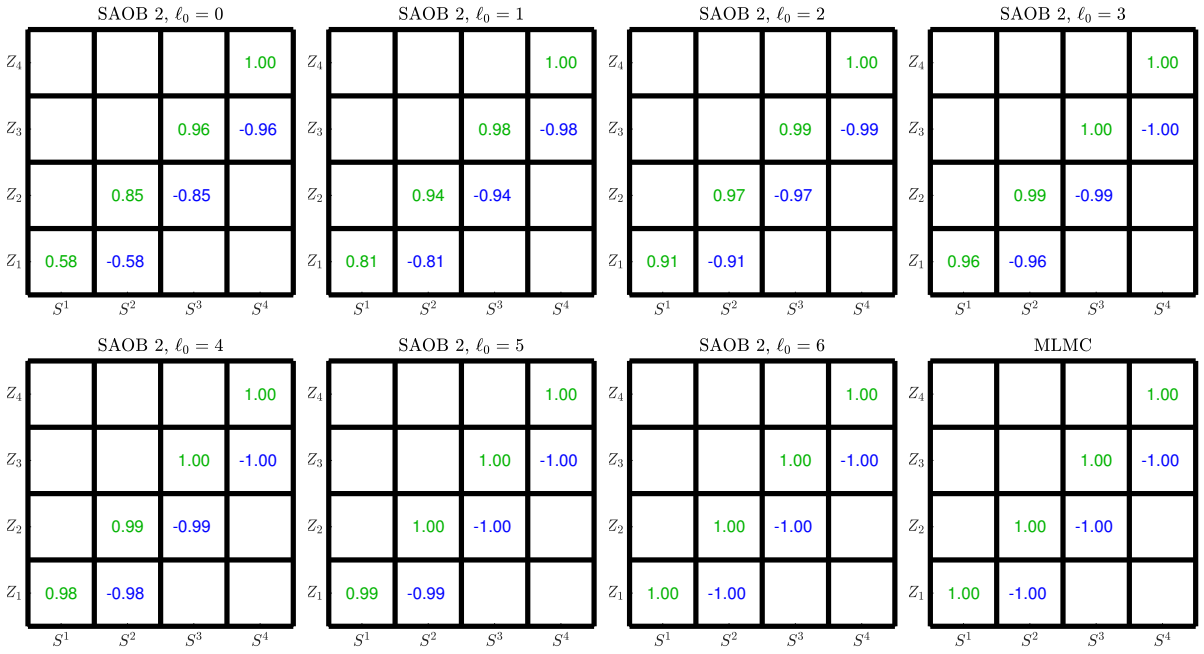

Figure 6.2: Coefficients $\beta^{k,\mathrm{SAOB}\,2}(\ell_0)$ of the SAOB 2 and $\beta^{k,\mathrm{MLMC}}$ of the MLMC estimator for different initial accuracies ℓ_0. The latter do not depend on ℓ_0, thus these coefficients are only drawn once in the bottom right. The error in terms of (6.47) is drawn in the left plot of Figure 6.1 in d_2.

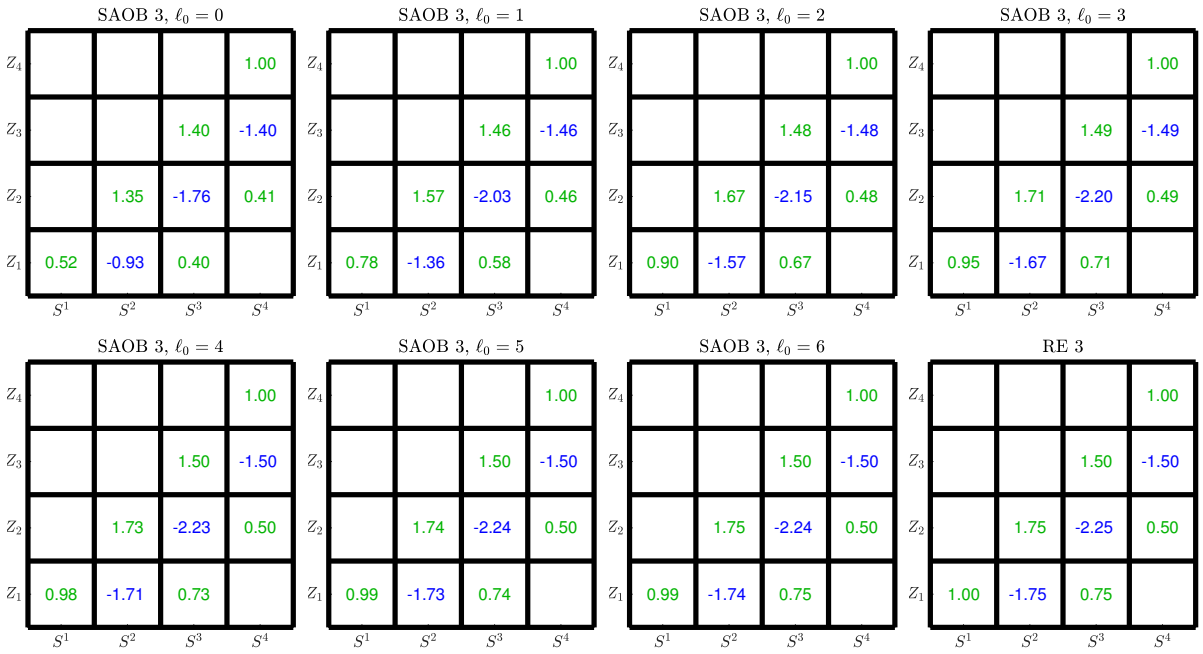

Figure 6.3: Coefficients $\beta^{k,\mathrm{SAOB}\,3}(\ell_0)$ of the SAOB 3 and $\beta^{k,\mathrm{RE}\,3}$ RE 3 estimator for different initial accuracies ℓ_0. The latter do not depend on ℓ_0, thus these coefficients are only drawn once in the bottom right. The error in terms of (6.47) is drawn in the left plot of Figure 6.1 in d_3.

same model groups as the RE κ estimator. Although this result seems reasonable due to the geometric cost increase, it is far from obvious and clearly depends on the involved constants. Once such a result is established, it should be straightforward to verify the convergence of the coefficients. We outline this step for SAOB 4. First, the bias constraint requires us to choose

$$\beta_4^4 = 1.$$

We now have three degrees of freedom left, namely β_1^4, β_2^4 and β_3^4. This allows to remove Z, c_2, c_3 from (6.46) and thus we obtain a variance rate of $\gamma_{\mathrm{Var}} = 2 \cdot 3 = 6$. This is the only linear combination independent of ℓ_0 to achieve this rate, all other linear combinations achieve at most the rate 4. Since the resulting estimator is now biased, we have to correct it using model groups with coarser models. We have to satisfy $\beta_{[3]}^3 = -\beta_{[3]}^4$ due to the bias constraint and thus we now only have two degrees of freedom left. Therefore only a single ℓ_0 independent unique linear combination leads to the rate 4. This process is now repeated for all model groups leading to unique coefficients, which are in fact the coefficients of the RE 4 estimator. Stated differently, the cost of the SAOB 4 is

$$\mathbb{W}\big[\widehat{\mu}_L^{\mathrm{SAOB}\,4}\big] = \left(\sum_{k=1}^{4}((\beta^k)^T C^k \beta^k W^k)^{1/2}\right)^2 \approx c\left(\sum_{k=1}^{4} 2^{-\gamma_k(k+\ell_0)+k}\right)^2,$$

which is minimal only if the rates γ_k are minimal. Thus, the SAOB 4 is forced to find the smallest rates which for coefficients independent of ℓ_0 is only achieved for the RE vectors. Since we are in the preasymptotic regime, the coefficient β^1 of the SAOB 4 of the coarsest model group may differ in the order of $2^{-\gamma_2 \ell_0} = 2^{-\ell_0}$ without worsening the asymptotic variance behaviour. This together with the fact that SAOB aims to keep the sum of the independent variables $\xi_1, \dots, \xi_4$ small explains the observed rate in the left plot of Figure 6.1. ◇

Improved bounds for SAOB. The RE estimator is used as upper bound for the SAOB, however, we now give an example where a different estimator achieves a larger variance reduction than the RE estimator. The basic idea to construct models such that the RE estimator is not able to remove terms in the expansion if the order is not polynomial. Let us illustrate this for the following QoI

$$Z_\ell(\omega) := Z(\omega) + c_2(\omega)(\ell+\ell_0)^2 2^{-(\ell+\ell_0)/2} + c_3(\omega)2^{-(\ell+\ell_0)} + \xi_\ell(\omega)2^{-3(\ell+\ell_0)/2}.$$

We assume that $Z, c_2, c_3, \xi_1, \xi_2, \cdots \sim N(0,1)$ are independent random variables. The RE 2 estimator is not able to remove the term $\ell^2 2^{-\ell/2}$ and does not improve any further. However, the SAOB is able to remove this term. To see why, we w.l.o.g. set $\ell_0 = 0$ and write down the linear combination

$$\sum_{\ell=1}^4 \beta_\ell Z_\ell = \left(\sum_{\ell=1}^4 \beta_\ell\right) Z + \left(\sum_{\ell=1}^4 \beta_\ell \ell^2 2^{-\ell/2}\right) c_2 + \left(\sum_{\ell=1}^4 \beta_\ell 2^{-\ell}\right) c_3 + \sum_{\ell=1}^4 \beta_\ell \xi_\ell 2^{-3\ell/2}.$$

We want that the coefficients β_ℓ satisfy three conditions

$$\sum_{\ell=1}^4 \beta_\ell = 0, \quad \sum_{\ell=1}^4 \beta_\ell 2^{-\ell} = 0, \quad \sum_{\ell=1}^4 \beta_\ell \ell^{-1} 2^{-2\ell} = 0$$

to remove the first three terms such that only the remainder with the higher order $2^{-3\ell/2}$ remains. We further want to ensure that

$$\beta_L = 1$$

to not obtain $\beta = 0$. This is now a 4×4 system that we are able solve. Therefore, if the $|\beta_\ell|$ are bounded by a constant independently of ℓ, we obtain the improved rate of $2^{-3\ell/2}$. We call this specific coefficient β^{LGS} and compare the variance of the linear combination

$$\mathbb{V}\left[\sum_{\ell=1}^4 \beta_\ell^{\mathrm{LGS}} Z_\ell\right] = (\beta^{\mathrm{LGS}})^T C \beta^{\mathrm{LGS}}$$

with the RE 2 and RE 3 difference vectors $v^{4,2} - v^{3,2}$ respectively $v^{4,3} - v^{3,3}$ assuming the (incorrect) rate $\gamma_2 = 1/2$. We further compare it to the variance obtained using the smallest eigenvector $u^{\min}$ of C in the following sense

$$\mathrm{Var}_{u^{\min}} = \mathbb{V}\left[\sum_{\ell=1}^4 \frac{u_\ell^{\min}}{u_4^{\min}} Z_\ell\right] = \frac{(u^{\min})^T C u^{\min}}{(u_4^{\min})^2}.$$

The division ensures that the coefficient in front of Z_4 is one, since $u_4^{\min} \neq 0$ for this experiment. In this case, we have $\mathbb{V}^{\min} = \mathrm{Var}_{u^{\min}}$. We plot the resulting difference to the eigenvalue and variance

$$\begin{aligned} e_u(\ell_0) &:= \|\beta^{\mathrm{LGS}}(\ell_0) - u^{\min}(\ell_0)/u_4^{\min}(\ell_0)\|, \\ e_{\mathrm{Var}}(\ell_0) &:= (u^{\min}(\ell_0))^T C(\ell_0) u^{\min}(\ell_0)/(u_4^{\min}(\ell_0))^2 - (\beta^{\mathrm{LGS}}(\ell_0))^T C(\ell_0) \beta^{\mathrm{LGS}}(\ell_0) \end{aligned}$$

in Figure 6.4, where we now make the dependence on ℓ_0 explicit. As predicted, the vector β^{LGS} is able to improve the rate such that only the remainder with a rate $\gamma_{\mathrm{Var}} = 2 \cdot 3/2 = 3$ remains. The RE vectors are not able to achieve this. We further conclude that the vector β^{LGS} converges to the scaled eigenvector $u^{\min}$ corresponding to the smallest eigenvalue and the respective variance of the linear combination converges to $\mathbb{V}^{\min}$.

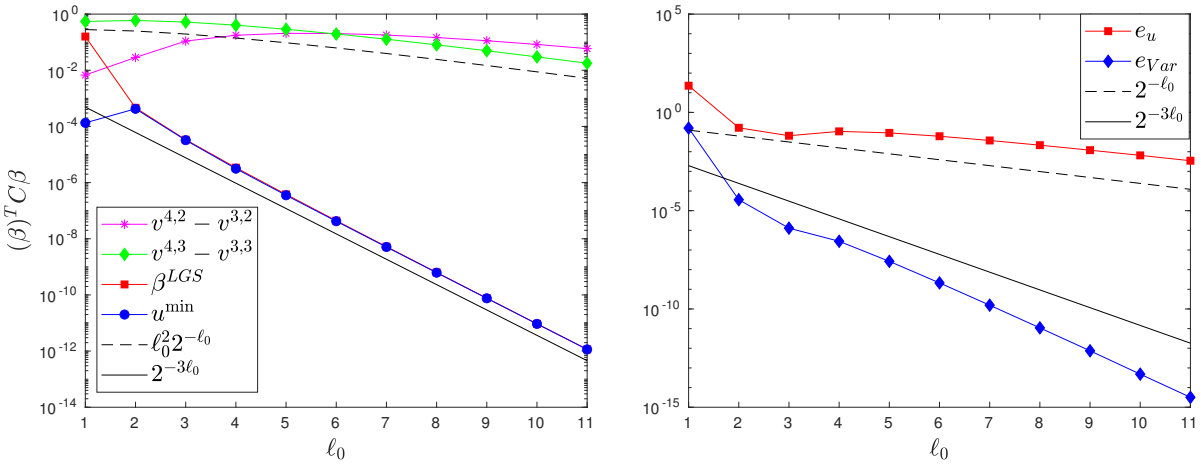

Figure 6.4: The left plot shows the variance of certain linear combinations where we replace β with the corresponding value. The right image shows the convergence of the variance and the coefficients of β^{LGS} to the scaled eigenvector $u^{\min}/u_4^{\min}$ for the smallest eigenvalue of C. Reference rates are drawn black.

Remark 6.37 (Generalized estimator)**.** We try to generalize the estimator from above. Assume that the QoI satisfies

$$Z_\ell := Z + \sum_{j=2}^{q-1} c_j \varphi_j(\ell) + r_\ell, \tag{6.49}$$

where r_ℓ is a remainder term decaying sufficiently fast. With $\varphi_1(\ell) := 1$ we look for coefficients β that satisfy

$$\begin{aligned} \sum_{\ell=1}^{q-1} \beta_\ell \varphi_j(\ell) = 0 \qquad & \text{for all } j = 1, \dots, q-1, \\ \beta_q = 1. & \end{aligned} \tag{6.50}$$

Under the assumption that this linear system has a solution the variance is

$$\mathbb{V}\left[\sum_{\ell=1}^{q} \beta_\ell Z_\ell\right] = \mathbb{V}\left[\sum_{\ell=1}^{q} \beta_\ell r_\ell\right],$$

that is only some linear combination of the remainder has to be estimated. Again, using the properties of the shift we are able to define consecutive vectors

$$y^{\ell,q} := D y^{\ell-1,q} \qquad \text{for all } \ell \geq q.$$

All that remains is to define the vectors $y^{\ell,q}$ for $\ell < q$ using the conditions (6.50), which is left as an exercise for the reader. Clearly, if (6.50) is satisfied, then we have a reduced variance. It is straightforward to define an estimator similar to the RE estimator in Definition 6.17

$$\widehat{\mu}_{y^{L,q}} = \sum_{\ell=1}^{L} \sum_{j=1}^{L} (y_j^{\ell,q} - y_j^{\ell-1,q}) \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} Z_j^{i,\ell}.$$

In fact, if we have $\varphi_j(\ell) = 2^{-\gamma_j \ell}$ in (6.50), then this estimator is precisely the RE estimator except for the scaling of the last coefficient $\beta_q = 1$. For other biases we may introduce weights $a_1, \dots, a_L$ similarly to (6.29). ◇

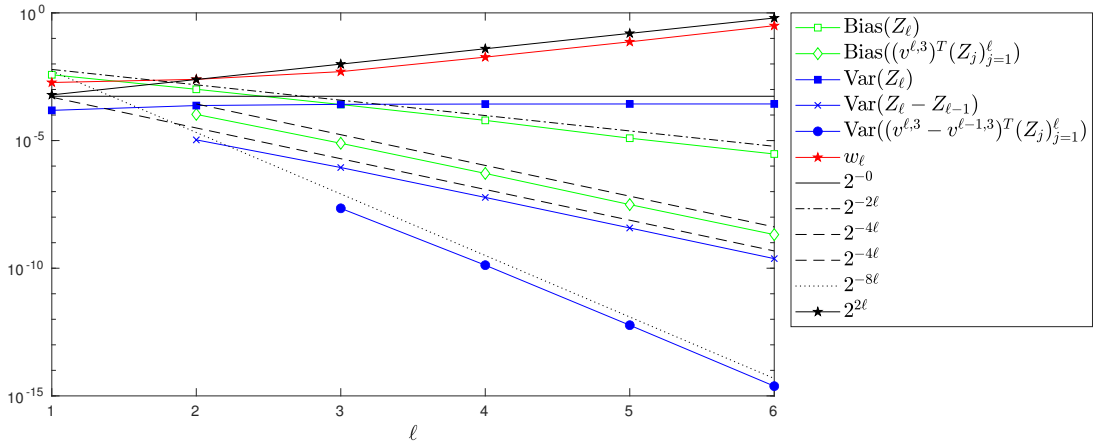

Figure 6.5: Biases, variances and costs for the QoI defined in (6.52). Here $Z_\ell - Z_{\ell-1}$ corresponds to the difference $(v^{\ell,2} - v^{\ell-1,2})^T(Z_j)_{j=1}^{\ell}$.

While it is possible to generalize the RE estimator to other estimators with improved variance reduction if an expansion of the form (6.49) is available and known a-priori, it is more involved due the more general φ_j. We never apply this for a practically relevant problem and thus we do not analyse this method any further.

6.5 Numerical experiments with an elliptic PDE

PDE example with true costs. We use the PDE example from Section 2.3 such that $\mathbb{P}$–almost surely

$$\begin{aligned} -\operatorname{div}(a(x,\omega)\nabla y(x,\omega)) &= 1, \quad x \in (0,1)^2, \\ y(x,\omega) &= 0, \quad x \in \partial(0,1)^2. \end{aligned} \tag{6.51}$$

The diffusion coefficient is lognormal $a := \exp(b)$, where b is a mean zero Gaussian random field on $(0,1)^2$ with Whittle–Matérn covariance with smoothness $\nu := 3/2$, variance $\sigma^2 := 1$ and correlation length $\ell := 0.5$. We integrate the weak solution y over $D_{\text{obs}} := (3/4, 7/8) \times (7/8, 1)$ to define the QoI

$$Z := \frac{1}{|D_{\text{obs}}|} \int_{D_{\text{obs}}} y(x)dx. \tag{6.52}$$

We discretize Z by discretizing y using linear finite elements with uniform mesh refinement and compute the corresponding covariance matrix C with 10^5 pilot samples. We exclude the cost of the pilot samples from the subsequent analysis in this section. Notice that the parameters in (6.51) are smooth in the sense that the diffusion a as well as the constant right–hand side 1 is smooth and we integrate over y in (6.52). This gives a smooth QoI such that Assumption 6.10 and Assumption 6.11 should be satisfied with coefficients $q = q_{\text{path}} = q_{\text{mean}} = 3$ and $\gamma_1 = 0$, $\gamma_2 = 2$ and $\gamma_3 = 4$. Since we solve the PDE in 2 dimensions, we expect a doubling of the cost with uniform mesh refinement and a rate of $\gamma_{\text{Cost}} = 2$. We plot the biases, variances and the costs in Figure 6.5 to verify these claims numerically. This result shows that we are able to use RE to potentially obtain improved variance reduction. The asymptotic costs of every estimator discussed in this thesis are

Estimator	κ	γ_{Bias}	γ_{Var}	Complexity bound	HF evaluations	Source
$\widehat{\mu}_L^{\text{MC}}$	1	2	0	$\varepsilon^{-1} + \varepsilon^{-3}$	ε^{-2}	Theorem 3.11
$\widehat{\mu}_L^{\text{MLMC}}$	2	2	4	$\varepsilon^{-1} + \varepsilon^{-2}$	$\varepsilon^{-0.5}$	Theorem 3.49
$\widehat{\mu}_L^{\text{MFMC}}$	L	2	4	$\varepsilon^{-1} + \varepsilon^{-2}$	$\varepsilon^{-0.5}$	Theorem 3.31
$\widehat{\mu}_L^{\text{SAOB 2}}$	2	2	4	$\varepsilon^{-1} + \varepsilon^{-2}$	$\varepsilon^{-0.5}$	Corollary 6.26
$\widehat{\mu}_L^{\text{SAOB 3}}$	3	2	8	$\varepsilon^{-1} + \varepsilon^{-2}$	$\varepsilon^{0.5}$	Corollary 6.26
$\widehat{\mu}_L^{\text{SAOB}}$	L	2	8	$\varepsilon^{-1} + \varepsilon^{-2}$	$\varepsilon^{0.5}$	Corollary 6.26
$\widehat{\mu}_{v^{L,3}}^{\text{MC}}$	1	4	0	$\varepsilon^{-0.5} + \varepsilon^{-2.5}$	ε^{-2}	Example 6.6
$\widehat{\mu}_{v^{L,3}}^{\text{RE 2}}$	2	4	4	$\varepsilon^{-0.5} + \varepsilon^{-2}$	$\varepsilon^{-1.25}$	Theorem 6.25
$\widehat{\mu}_{v^{L,3}}^{\text{RE 3}}$	3	4	8	$\varepsilon^{-0.5} + \varepsilon^{-2}$	$\varepsilon^{-0.75}$	Theorem 6.20
$\widehat{\mu}_{v^{L,3}}^{\text{SAOB 2}}$	2	4	4	$\varepsilon^{-0.5} + \varepsilon^{-2}$	$\varepsilon^{-1.25}$	Corollary 6.26
$\widehat{\mu}_{v^{L,3}}^{\text{SAOB 3}}$	3	4	8	$\varepsilon^{-0.5} + \varepsilon^{-2}$	$\varepsilon^{-0.75}$	Corollary 6.21
$\widehat{\mu}_{v^{L,3}}^{\text{SAOB}}$	L	4	8	$\varepsilon^{-0.5} + \varepsilon^{-2}$	$\varepsilon^{-0.75}$	Corollary 6.26

Table 6.1: Different estimators with the respective coupling number κ that denotes the maximum number of models using the same event ω, their bias and variance reduction rate as well as an upper bound on their cost complexity (6.53) for the true cost rate $\gamma_{\text{Cost}} = 2$. The first term always corresponds to the ceiling cost $\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}$, whereas the second value corresponds to the variance and second part of (6.53). The column "HF evaluations" contains the upper bound on the number of high fidelity evaluations, which can be derived from Corollary 6.2. The last column contains the references and proofs for the upper complexity bound.

of the form (6.9)

$$\mathbb{W}[\widehat{\mu}_L] \leq c\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} + c\begin{cases} \varepsilon^{-2}, & \text{if } \gamma_{\text{Var}} > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } \gamma_{\text{Var}} = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-\gamma_{\text{Var}}}{\gamma_{\text{Bias}}}}, & \text{if } \gamma_{\text{Var}} < \gamma_{\text{Cost}} \end{cases} \tag{6.53}$$

for different values of γ_{Bias} and γ_{Var}. We write down the cost of the estimators in Table 6.1 and describe how we derived these values. First of all, for all estimators with $\alpha = e_L$ the bias rate $\gamma_{\text{Bias}} = 2$ and for estimators with $\alpha = v^{L,3}$ the rate $\gamma_{\text{Bias}} = 4$ holds. The variance reduction rate should be $\gamma_{\text{Var}} = 2\gamma_\kappa$ for an optimal variance reduction, where κ is the coupling number. This however, is in general not reached, since either the QoI is not smooth enough, thus $q_{\text{path}} < \kappa$ or the coefficients of the estimator are chosen sub optimally. Applying this logic, this means that the SAOB κ with $\kappa = 2$ respectively $\kappa = 3$ reaches $\gamma_{\text{Var}} = 4$ respectively $\gamma_{\text{Var}} = 8$. This also holds for the RE estimators. The SAOB with no coupling restriction $\kappa = L$ may reach a larger variance reduction rate and thus $\gamma_{\text{Var}} = 8$ is a lower bound. The MC estimator does not have any variance reduction $\gamma_{\text{Var}} = 0$. The previous explanations concerning the variance are valid for all estimators except for the MFMC estimator, which couples all L models but achieves a smaller variance reduction rate due to suboptimal coefficients. We later give a heuristic reason for this. We compute the upper complexity bound with these values using (6.53). For a rigorous proof, we refer to the statement in the last column of Table 6.1.
We plot the cost of every estimator to reach a MSE of ε^2 in Figure 6.6 for both $\alpha = e_\ell$ and using RE once with $\alpha = v^{\ell,3}$. We conclude that the upper bounds for the complexity in Table 6.1 predict the correct values.
We now examine the number of high fidelity model evaluations, where the upper bounds

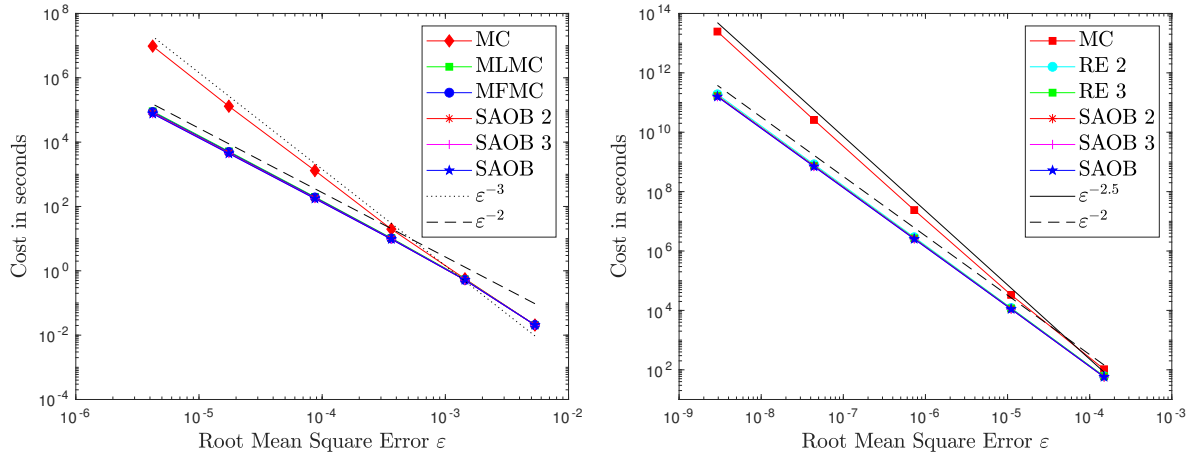

Figure 6.6: Complexity of different estimators for different RMSE ε for the true cost rate $\gamma_{\text{Cost}} = 2$. The left image shows estimators with bias e_ℓ and the right image with $v^{\ell,3}$, i.e. for SAOB the left plot shows $\widehat{\mu}_L^{\text{SAOB}}$ and the right one shows $\widehat{\mu}_{v^{\ell,3}}^{\text{SAOB}}$.

for $k = L$ are given in Table 6.1. The actual numbers used by the estimators are in Figure 6.7 and they match the predictions. For $\alpha = v^{\ell,3}$ the number of high fidelity evaluations does not decrease to zero, which allows us to use more and more pilot samples for the sample covariance matrix without worsening the asymptotic complexity.

PDE example with artificial costs. All estimators, except for standard MC, achieve the optimal complexity of ε^{-2} with the true cost rate $\gamma_{\text{Cost}} = 2$. Therefore, we artificially increase this rate to $\gamma_{\text{Cost}} = 6$ by setting $w_\ell = 10^{-6} \cdot 2^{6\ell}$. Similarly to before, we derive Table 6.2 containing the updated complexity bound and the bound on the number of high fidelity evaluations. We again plot the computational complexity in Figure 6.8 and the number of high fidelity evaluations in Figure 6.9. We conclude that changing the bias from e_ℓ to $v^{\ell,3}$ is necessary to reduce the complexity of the estimation, since otherwise we have at least costs of ε^{-3} due to the ceiling of the number of samples. This ceiling cost has to be reduced by increasing the bias rate from 2 to 4 by using RE. It is also apparent that the number of high fidelity evaluations tends to zero if we use more than two models for $\alpha = e_\ell$, i.e. the RE 3 or SAOB 3 estimator. This is typically not desirable and indicates that the bias vector can and should be chosen differently to reduce the overall complexity. The bias vector $\alpha = v^{\ell,3}$ yields the best possible complexity of ε^{-2} for estimators that couple at least three models except for MFMC, i.e. the SAOB κ and the weighted RE κ estimator for $\kappa \geq 3$.

Coefficients for ACV estimators and the MFMC estimator. We comment on the ACV estimators, which we did not include up until now. This is deliberate, since the asymptotic cost of the ACV estimators is not known to date. For the same QoI from (6.52) with artificial cost rate $\gamma_{\text{Cost}} = 6$ we plot the cost of these estimators in Figure 6.10. The ACV estimators have the same asymptotic complexity as the MLMC estimator and do not reach the improved asymptotics of the RE estimators, even if we exclude ceiling the number of samples. This also holds for the MFMC estimator and is somewhat surprising, since these estimators couple more models than MLMC. We provide a possible explanation in terms of their respective coefficients. As we have already seen in Section 4.5 and which can be verified quite easily, the coefficients of all ACV estimators

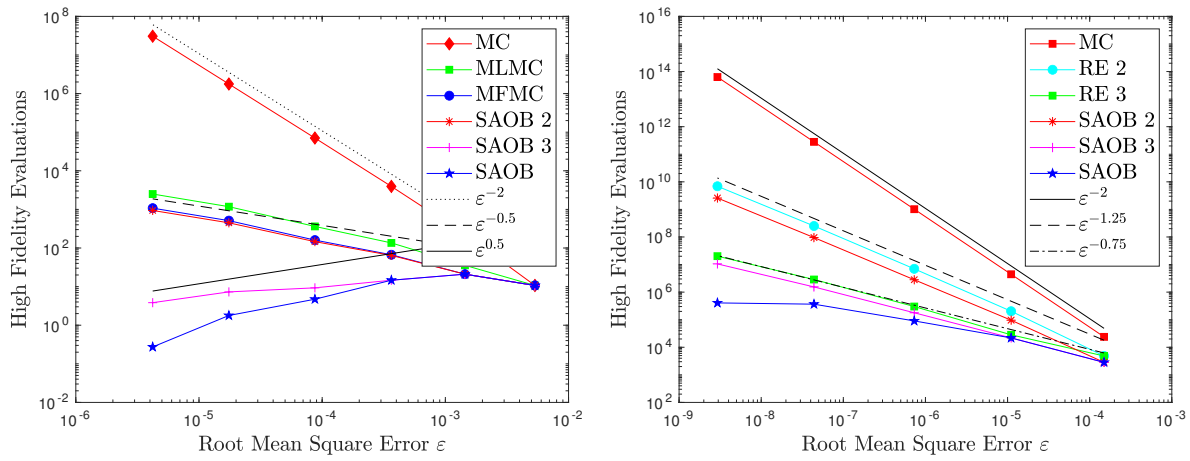

Figure 6.7: Number of high fidelity evaluations for different RMSE ε for the true cost rate $\gamma_{\text{Cost}} = 2$. The left image shows estimators with bias $\alpha = e_\ell$ and the right image shows the bias $v^{\ell,3}$, i.e. for SAOB the left plot shows $\widehat{\mu}_L^{\text{SAOB}}$ and the right one shows $\widehat{\mu}_{v^{\ell,3}}^{\text{SAOB}}$.

Estimator	κ	γ_{Bias}	γ_{Var}	Complexity bound	HF eval bound	Source
$\widehat{\mu}_L^{\text{MC}}$	1	2	0	$\varepsilon^{-3} + \varepsilon^{-5}$	ε^{-2}	Theorem 3.11
$\widehat{\mu}_L^{\text{MLMC}}$	2	2	4	$\varepsilon^{-3} + \varepsilon^{-3}$	ε^{0}	Theorem 3.49
$\widehat{\mu}_L^{\text{MFMC}}$	L	2	4	$\varepsilon^{-3} + \varepsilon^{-3}$	ε^{0}	Theorem 3.31
$\widehat{\mu}_L^{\text{SAOB 2}}$	2	2	4	$\varepsilon^{-3} + \varepsilon^{-3}$	ε^{0}	Corollary 6.26
$\widehat{\mu}_L^{\text{SAOB 3}}$	3	2	8	$\varepsilon^{-3} + \varepsilon^{-2}$	$\varepsilon^{0.5}$	Corollary 6.26
$\widehat{\mu}_L^{\text{SAOB}}$	L	2	8	$\varepsilon^{-3} + \varepsilon^{-2}$	$\varepsilon^{0.5}$	Corollary 6.26
$\widehat{\mu}_{v^{L,3}}^{\text{MC}}$	1	4	0	$\varepsilon^{-1.5} + \varepsilon^{-3.5}$	ε^{-2}	Example 6.6
$\widehat{\mu}_{v^{L,3}}^{\text{RE 2}}$	2	4	4	$\varepsilon^{-1.5} + \varepsilon^{-2.5}$	ε^{-1}	Theorem 6.25
$\widehat{\mu}_{v^{L,3}}^{\text{RE 3}}$	3	4	8	$\varepsilon^{-1.5} + \varepsilon^{-2}$	$\varepsilon^{-0.25}$	Theorem 6.20
$\widehat{\mu}_{v^{L,3}}^{\text{SAOB 2}}$	2	4	4	$\varepsilon^{-1.5} + \varepsilon^{-2.5}$	ε^{-1}	Corollary 6.26
$\widehat{\mu}_{v^{L,3}}^{\text{SAOB 3}}$	3	4	8	$\varepsilon^{-1.5} + \varepsilon^{-2}$	$\varepsilon^{-0.25}$	Corollary 6.21
$\widehat{\mu}_{v^{L,3}}^{\text{SAOB}}$	L	4	8	$\varepsilon^{-1.5} + \varepsilon^{-2}$	$\varepsilon^{-0.25}$	Corollary 6.26

Table 6.2: Different estimators with the respective coupling number κ that denotes the maximum number of models using the same event ω, their bias and variance reduction rate as well as an upper bound on their cost complexity (6.53) for the artificial cost rate $\gamma_{\text{Cost}} = 6$. The first term always corresponds to the ceiling cost $\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}$, whereas the second value corresponds to the variance and second part of (6.53). The column "HF eval bound" contains the upper bound on the number of high fidelity evaluations, which can be derived from Corollary 6.2. The last column contains the references and proofs for the upper complexity bound.

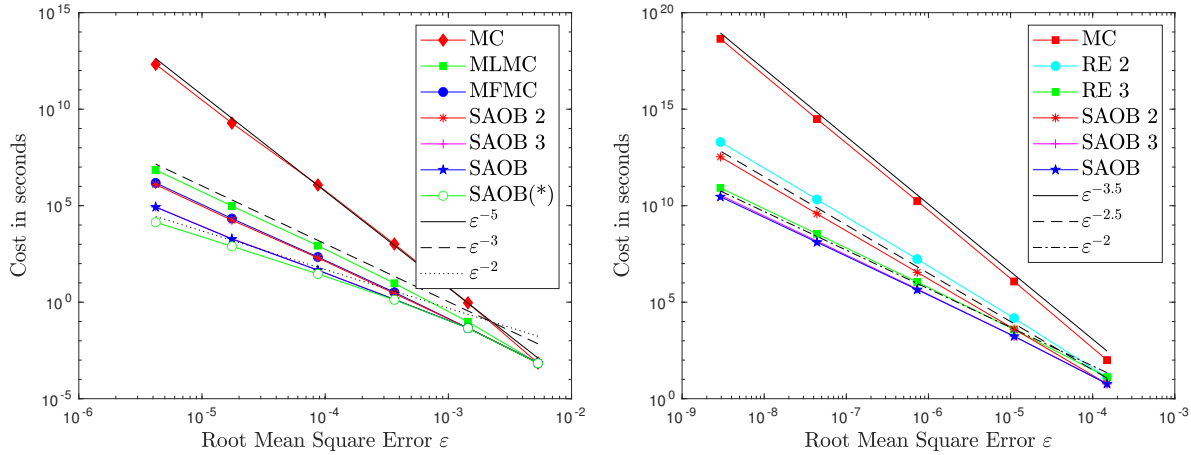

Figure 6.8: Complexity of different estimators for different RMSE ε for the artificial cost rate $\gamma_{\text{Cost}} = 6$. The left image shows estimators with bias e_ℓ and the right image with $v^{\ell,3}$, i.e. for SAOB the left plot shows $\widehat{\mu}_L^{\text{SAOB}}$ and the right one shows $\widehat{\mu}_{v^{\ell,3}}^{\text{SAOB}}$. Here SAOB(*) is the cost of the SAOB without ceiling the number of samples.

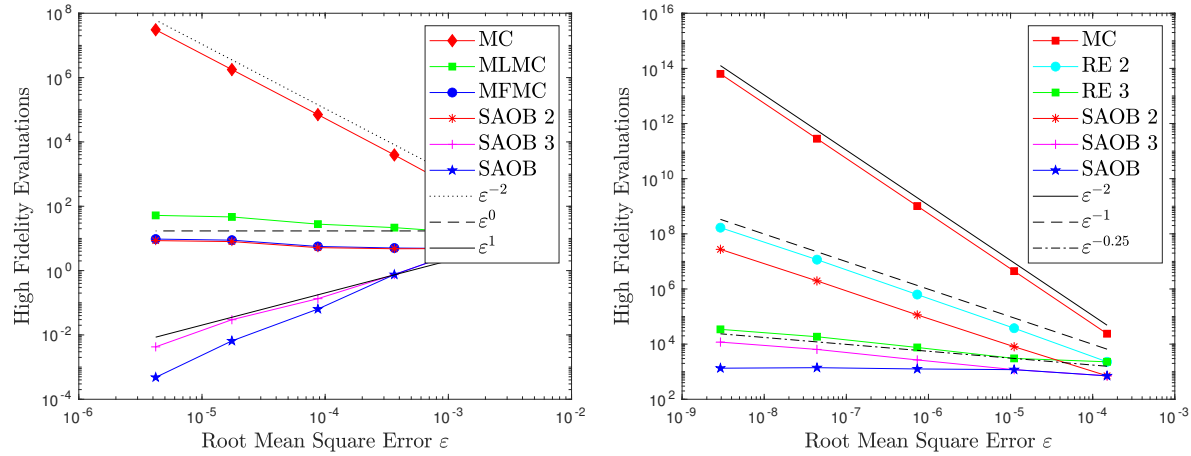

Figure 6.9: Number of high fidelity evaluations for different RMSE ε for the artificial cost rate $\gamma_{\text{Cost}} = 6$. The left image shows estimators with bias e_ℓ and the right image shows the bias $v^{\ell,3}$, i.e. for SAOB the left plot shows $\widehat{\mu}_L^{\text{SAOB}}$ and the right one shows $\widehat{\mu}_{v^{\ell,3}}^{\text{SAOB}}$.

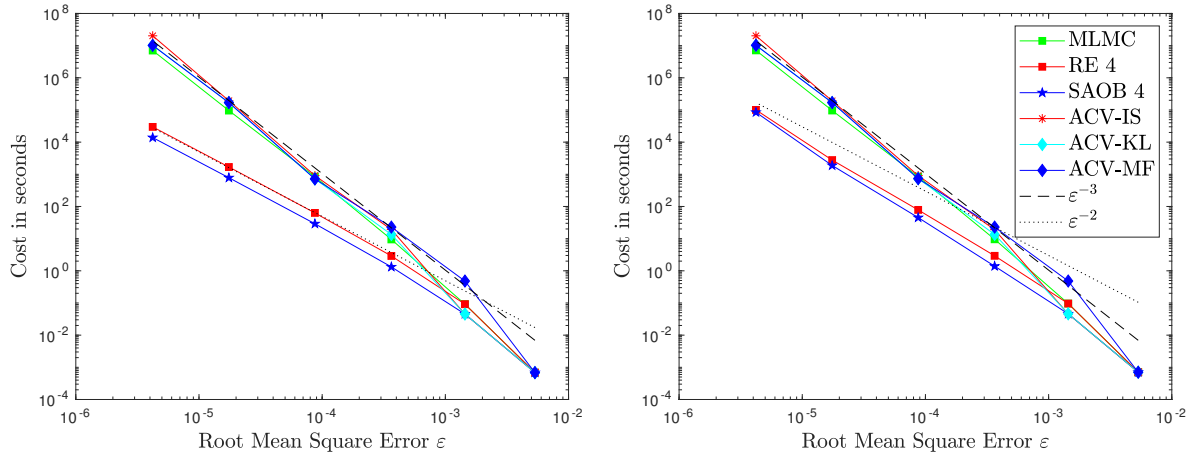

Figure 6.10: Complexity of different estimators for different RMSE ε for the artificial cost rate $\gamma_{\text{Cost}} = 6$. The left image shows the complexity of the estimators without ceiling and the right image with ceiling. Both images show estimators with a bias of $\alpha = e_\ell$. In particular, the RE 4 estimator is $\widehat{\mu}_L^{\text{RE}\,4}$.

satisfy the sign condition

$$\operatorname{sign}(\beta_\ell^L) = -\operatorname{sign}(\beta_\ell^k) \qquad \text{for all } k \in \{1, \dots, L-1\} \text{ with } \ell \in S^k. \tag{6.54}$$

This simply means that the sign of the coefficients can be deduced from the sign of the coefficients of the model group that uses all models $\{1, \dots, L\}$. Similarly, the coefficients of the MFMC estimator also satisfy a sign condition

$$\operatorname{sign}(\beta_\ell^\ell) = -\operatorname{sign}(\beta_\ell^k) \qquad \text{for all } k \in \{\ell+1, \dots, L\} \text{ with } \ell \in S^k, \tag{6.55}$$

thus the sign is given in terms of the diagonal element. However, looking at the difference of two consecutive RE vectors (6.21) with $L > q$ shows that

$$\begin{aligned}\Delta v^q = D v^{q-1,q} - v^{q-1,q} = \frac{1}{\prod_{j=2}^{q-1}(2^{\gamma_j}-1)} \begin{pmatrix} 0 \\ (-1)^q \\ (-1)^{q-1} \sum_{2 \le i_1 \le q-1} 2^{\gamma_{i_1}} \\ (-1)^{q-2} \sum_{2 \le i_1 < i_2 \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}} \\ \vdots \\ \sum_{2 \le i_1 < i_2 < \cdots < i_{q-3} \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\cdots+\gamma_{i_{q-3}}} \\ -\sum_{2 \le i_1 < i_2 < \cdots < i_{q-2} \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\cdots+\gamma_{i_{q-2}}} \\ \sum_{2 \le i_1 < i_2 < \cdots < i_{q-1} \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\cdots+\gamma_{i_{q-1}}} \end{pmatrix} \\ - \frac{1}{\prod_{j=2}^{q-1}(2^{\gamma_j}-1)} \begin{pmatrix} (-1)^q \\ (-1)^{q-1} \sum_{2 \le i_1 \le q-1} 2^{\gamma_{i_1}} \\ (-1)^{q-2} \sum_{2 \le i_1 < i_2 \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}} \\ \vdots \\ \sum_{2 \le i_1 < i_2 < \cdots < i_{q-3} \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\cdots+\gamma_{i_{q-3}}} \\ -\sum_{2 \le i_1 < i_2 < \cdots < i_{q-2} \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\cdots+\gamma_{i_{q-2}}} \\ \sum_{2 \le i_1 < i_2 < \cdots < i_{q-1} \le q-1} 2^{\gamma_{i_1}+\gamma_{i_2}+\cdots+\gamma_{i_{q-1}}} \\ 0 \end{pmatrix} \in \mathbb{R}^q.\end{aligned}$$

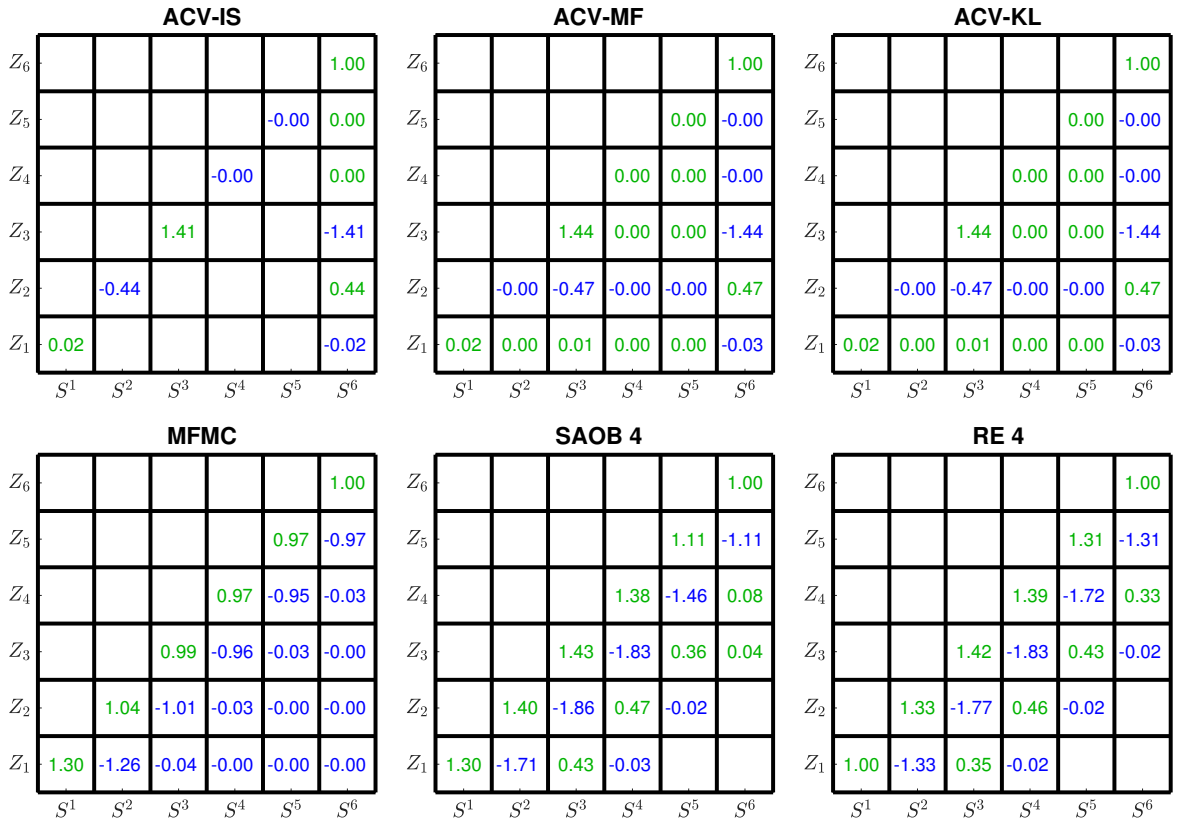

Figure 6.11: Coefficients for different estimator for a RMSE $\varepsilon = 3 \cdot 10^{-6}$ for the artificial cost rate $\gamma_{\text{Cost}} = 6$. Here the coefficients of the ACV–MF and ACV–KL estimator are the same. The RE estimator exhibits a checkerboard pattern, whereas the SAOB 4 also shows this pattern except for a small deviation for S^6 and Z_3.

Taking a look at the sign of these vectors shows that both $Dv^{q-1,q}$ and $-v^{q-1,q}$ have entries with equal signs. We use that $\gamma_j \geq 0$ for $j \in \{1, \dots, q-1\}$ to conclude

$$\operatorname{sign}((\Delta v^q)_\ell) = \operatorname{sign}(v_{\ell-1}^{q-1,q}) = (-1)^{q-\ell} \qquad \text{for all } \ell \in \{2, \dots, q\}$$

and the special case of $\operatorname{sign}((\Delta v^q)_1) = (-1)^{q-1}$. Then

$$\operatorname{sign}((\Delta v^q)_\ell) = (-1)^{q-\ell} \qquad \text{for all } \ell \in \{1, \dots, q\}.$$

A further computation in the same spirit shows that the same is true independently of q

$$\operatorname{sign}((\Delta v^j)_\ell) = (-1)^{j-\ell} \qquad \text{for all } \ell \in \{1, \dots, j\} \quad \text{for all } j \in \{1, \dots, L\}. \tag{6.56}$$

Hence the coefficients of the RE estimators form a checkerboard pattern w.r.t. the sign of the linear combination. This however, cannot be achieved by either of the ACV or MFMC estimators due to (6.54) or (6.55). We verify these claims by plotting the coefficients of the ACV estimators along with the coefficients of the MFMC estimator, the SAOB 4 and the RE 4 estimator in Figure 6.11 for the smallest RMSE of $\varepsilon = 3 \cdot 10^{-6}$. We conclude that the conditions (6.54), (6.55) and (6.56) hold for this example. In general, if the coefficients of the RE estimators are mostly unique and often the only sensible choice to achieve a smaller variance, then this means that the ACV and MFMC estimators are not able to reach a higher variance reduction than MLMC due to the sign restriction. Therefore, even though these estimators couple more models than the MLMC estimator,

they do not improve due to their design, which restricts the number of possible linear combinations.

Remark 6.38 (Other models)**.** Assumption 6.10 and Assumption 6.11 might not be true for other model hierarchies and thus an improved variance reduction might be obtained even if condition (6.54) for ACV and condition (6.55) for MFMC holds. In any case, it is in general feasible to obtain complexity expression for the MLMC and the RE estimator since their coefficients β^k are fixed and known, which is not the case for both MFMC and ACV. This makes it difficult to derive statements about their complexity. The coefficients β^k of the SAOB and SAOB κ are also given implicitly, however for these estimators we have the principle that they are optimal in the class of linear unbiased estimators Theorem 6.3, which allows us to at least derive some upper bounds. ◇

Chapter 7

A multilevel approach for the risk neutral optimal control problem

In this chapter we examine the risk neutral optimal control problem. The goal is to steer the result of a system to a prescribed desired state. The system, described by an elliptic PDE, contains random functions and is controlled by a deterministic distributed control without control constraints. We account for the randomness by ensuring that the expected value of the distance to the desired state is small. The use and analysis of the MC method to solve this problem is more or less straightforward, however, the resulting costs to compute an accurate approximation are large. A straightforward application of MLMC to reduce the asymptotic complexity often leads to an ill–posed minimization problem. We overcome this difficulty by proposing a strategy where we apply the MLMC estimator to the deterministic control as opposed to the cost function where the expectation appears. The result is an estimator that has the improved complexity of MLMC without the drawbacks of being ill–posed. We divide this chapter into the following sections:

- In Section 7.1 we introduce the risk neutral optimal control problem. In essence, this is an infinite dimensional quadratic optimization problem with a linear PDE constraint. The cost function contains randomness, which we eliminate by taking the mean. We show existence and uniqueness of a solution, we derive the optimality conditions, give a short literature review and derive convergence results for the spatial discretization with finite elements. These results are standard and the focus of this chapter is the discretization of the mean, which we focus on in the remaining sections of this chapter.

- We continue with the standard MC approach in Section 7.2, which we use as a reference. This approach allows us to discuss the basic concepts needed for the MLMC approach. We show that the MC discretization leads to a well–posed problem and derive an asymptotic complexity result. The resulting costs of this method are typically quite large.

- We improve the MC approach with an MLMC approach in Section 7.3. We call the method Multilevel Monte Carlo for the control (MLC). The idea is to compute the difference of the optimal controls on two consecutive levels, which itself is a solution of an optimal control problem. This approach achieves an MLMC complexity result similar to Theorem 3.49. In particular, the MLC improves the standard MC method and is well posed in contrast to directly applying the MLMC telescoping sum idea to the cost function.

- We conduct numerical experiments in Section 7.4 to verify the complexity results for MC and MLC. We verify that we have substantial gains compared to the MC estimator in the order of two magnitudes.

7.1 The risk neutral optimal control problem

Introduction. We are interested in the risk neutral optimal control problem

$$\begin{aligned} \min_{u\in U} \quad & J(u) := \frac{1}{2}\mathbb{E}\big[\|y(u)-y_d\|_Z^2\big] + \frac{\alpha}{2}\|u\|_U^2 \\ \text{such that } \mathbb{P}\text{–a. s.} \quad & y(u,\omega) = S(\omega)u. \end{aligned} \tag{7.1}$$

This is a quadratic optimization problem with linear equality constraint. We call y the random state such that $y(\omega) \in Y$, $u \in U$ the deterministic control and $y_d \in Z$ the desired state. We assume for simplicity that the desired state is deterministic, however, similar results can be obtained for random desired states. The state and control are coupled by the state equation such that $\mathbb{P}$–almost surely

$$y(u,\omega) = S(\omega)u.$$

We assume that the spaces Z, Y, U are Hilbert spaces with scalar product $(\cdot,\cdot)$ and induced norm $\|\cdot\|$. We drop the subscript in the norm and scalar product, since this will be clear from the surrounding context. We also demand that $Y \subseteq Z$ such that the image of $S(\omega)$ may strictly be smaller than Z. The goal is to approximate the optimal u that minimizes (7.1). In particular, the state y should be close to the desired state y_d on average in terms of the realizations of the linear solution operator $S(\omega) : U \to Y$. We furthermore use a Tikhonov regularization parameter $\alpha > 0$ since (7.1) is in general ill–posed for infinite dimensional problems or for finite dimensional problems if S has zero eigenvalues since multiple optimizers may exist. The case of not using a regularization term $\alpha = 0$ can be handled by control constraints, for example $a \le u \le b$ for real–valued $a, b \in \mathbb{R}$ with $a \le b$. This however, increases the technical difficulty of the analysis and the numerical experiments contrasting the goal of this chapter to introduce a variance reduction method for optimal control problems. We provide the example we use latter for the numerical experiments.

Example 7.1 (An optimal control problem)**.** We give the standard example also seen in [73, Section 1.5.3.1] without control constraints and with a random state equation. We have $Z := U := L^2(D)$ and $Y := H_0^1(D)$. The linear solution operator $S(\omega) : L^2(D) \to H_0^1(D)$ maps the distributed control u to the state y using the weak elliptic PDE with zero Dirichlet boundary conditions (2.9). We demand that $\mathbb{P}$–almost surely for all $v \in H_0^1(D)$

$$(a(\omega)\nabla y(\omega), \nabla v)_{L^2(D)} = (u, v)_{L^2(D)}. \tag{7.2}$$

The randomness in the diffusion coefficient a carries over to the solution operator S. We abbreviate condition (7.2) in short operator form as $y(u,\omega) = S(\omega)u$. The stochastic state then satisfies $y \in L^2(\Omega, Y)$, see Theorem 2.34. ◇

There is a lot of literature on optimization with deterministic PDE constraints [14, 15, 39, 73, 136]. The authors of [16] discuss how to introduce uncertainty for optimal control problems. More difficult PDE constraints with randomness compared to the elliptic case that we study was also done in [16] for a non–linear reaction–diffusion model and in [65] for the Navier–Stokes system. Shape optimization problems with random right hand side are discussed in [36].
The stochastic pathwise optimal control solves (7.1) without the mean. The optimal control is then stochastic. This was examined in [16] and implemented in [121] with

stochastic collocation and the stochastic Galerkin method. A method to compute the mean of the stochastic control with MLMC was done in [4]. However, if the randomness is crucial for the system, then the resulting optimal control does not approximate the solution of (7.1). Nevertheless, this is approach can be used to obtain an insight on how the randomness effects the solution of (7.1).

The stochastic collocation method is used to approximate the mean operator $\mathbb{E}$ in [135]. The authors of [17] combine this method with a multigrid method and apply it to a non–linear parabolic optimal control problem. Another method [18] is to obtain pathwise optimal controls that are used to compute a proper orthogonal decomposition. This decomposition is used to obtain a low dimensional but accurate approximation of U and thus the infinite dimensional (7.1) is replaced by a low dimensional problem. A similar approach using a reduced basis method is analysed in [27]. The collocation method was also used as part of a trust region algorithm, a general optimization procedure, in [83, 84].

A stochastic gradient method with approximate line search was analysed in [93] for the standard MC discretization of the gradient. A multilevel version with detailed analysis was performed in [94]. The authors of [138] use the MLMC estimator to improve the numerical complexity to compute an approximation of the gradient. Under the additional assumption that the norm of the gradient converges linearly to zero, the authors show that the optimal asymptotic complexity of ε^{-2} to reach a gradient norm of ε is achieved. The same authors in [139] also apply the MG/OPT framework, which is a multilevel technique for optimization [91, 100]. Additional to the variance reduction the idea is to use a coarse grid for the search direction along which we optimize. This framework is combined with stochastic collocation in [82]. Another approach using stochastic gradient descent with MC, where the descent algorithm is coupled with the mesh size, can be found in [53]. Another well–known stochastic optimization method is SAGA [41], which was applied in [95] for the risk neutral optimal control problem. A variance reduction method for optimization that reuses information from previous iterations was studied in [32, 101].

Other approaches include a full discretization of both the probability space and the finite element space [75], use quasi–Monte Carlo [66] or a polynomial chaos expansion with low rank tensors [52]. Our approach for the optimal control problem is different from stochastic optimization methods. We apply a variant of MLMC for the optimality conditions of (7.1) avoiding some of the pitfalls of classical MLMC. In particular, applying MLMC for the expectation in (7.1) or directly for the optimality conditions leads to a potentially ill–posed and non–convex optimization problem. We push the telescoping sum idea to the deterministic control $u \in U$ to circumvent this problem while retaining the benefits of improved asymptotic complexity of MLMC.

Other risk measures, which we do not discuss in this chapter, are studied theoretically in [81]. The authors of [3] and [138] consider adding a variance term such that solutions with smaller variance are preferred thus ideally decreasing the probability for outliers. Other, more advanced risk measures like the conditional value at risk are studied in [85]. The authors of [86] study a MLMC approach for functions and include minimization problems. This approach is able to achieve the optimal complexity (up to logarithmic factors). However, the authors study problems where the parameter we want to minimize is in a compact, real–valued interval. Their approach requires the discretization of the cost function on the entire parameter interval, which is not feasible in the high dimensional setting and this is acknowledged by the authors. In particular, a naive application of this method requires the full approximation of the cost function, i.e. if the control is a finite element function on a grid with 1000 degrees of freedom then we have to evaluate the entire

function on a 1000 dimensional space. In our opinion, this approach for minimization seems suboptimal, since we discretize the cost function far away from the minimum thus adding additional and probably unnecessary costs.
The problem (7.1) uses the mean squared deviation from the desired state, which is a risk neutral risk measure since we do not penalize large deviations from the mean of the state y.

Existence, uniqueness and optimality conditions. We want to derive that (7.1) has a unique solution for $\alpha > 0$. This requires certain regularity for the solution operator S which should be bounded by a constant whose moments are also bounded.

Assumption 7.2 (S is regular)**.** The solution operator S and its adjoint S^* are pathwise bounded, that is $\mathbb{P}$–almost surely

$$\begin{aligned} \|S(\omega)u\| &\leq q(\omega)\|u\| \qquad \text{for all } u \in U, \\ \|S^*(\omega)y\| &\leq q(\omega)\|y\| \qquad \text{for all } y \in Y. \end{aligned}$$

The scalar valued random variable q has bounded fourth moment. S and S^* are random variables. ◇

We derive the necessary optimality conditions of (7.1). The directional derivative ∂_v of the direction $v \in U$ is

$$\begin{aligned} \partial_v \|u\|^2 &= \lim_{h\to 0} \frac{\|u+hv\|^2 - \|u\|^2}{h} = \lim_{h\to 0} \frac{(hv, 2u+hv)}{h} = 2(v,u), \\ \partial_v \mathbb{E}\big[\|y(u)-y_d\|^2\big] &= \lim_{h\to 0} \frac{\mathbb{E}[\|S(u+hv)-y_d\|^2] - \mathbb{E}[\|Su-y_d\|^2]}{h} \\ &= \lim_{h\to 0} \frac{\mathbb{E}[(Shv, 2(Su-y_d)+sv)]}{h} \\ &= 2\mathbb{E}[(Sv, Su-y_d)] = 2(v, \mathbb{E}[S^*Su - S^*y_d]). \end{aligned}$$

A necessary condition for optimality is $\partial_v J = 0$ for all $v \in U$ and thus

$$\alpha u + \mathbb{E}[S^*S]u = \mathbb{E}[S^*]y_d.$$

We rewrite this as operator equation to obtain the first order necessary optimality conditions,

$$(\alpha I + Q)u = f,$$

where we abbreviate $Q := \mathbb{E}[S^*S]$ and $f := \mathbb{E}[S^*]y_d$. We further define $B := \alpha I + Q$ to arrive at

$$Bu = f. \tag{7.3}$$

We now formally state the existence and uniqueness results of the optimal solution u of (7.1) which is the solution of (7.3). Hence this first order necessary optimality condition is also sufficient. This result is available in [27, 66, 75] and follows from standard arguments, e.g. in [73, Section 1.5.1].

Theorem 7.3 (Existence and uniqueness)**.** Let $\alpha > 0$ and Assumption 7.2 be true. Then there exists a unique minimizer $u \in U$ of (7.1) satisfying (7.3) and

$$\|u\| \leq c\|y_d\|. \tag{7.4}$$

Furthermore, the operator B is bounded, self–adjoint and invertible with bounded inverse

$$\|B^{-1}\| \leq c. \tag{7.5}$$

Proof. We verify that (7.3) has a unique solution using the Lax–Milgram Lemma [50, Section 6.2.1] for the bilinear form

$$a(u, v) = (Bu, v).$$

We verify that a is bounded w.r.t. both arguments with Assumption 7.2

$$\begin{aligned} |a(u,v)| &\leq \|\alpha u + \mathbb{E}[S^*S]u\|\|v\| \leq (\alpha\|u\| + \mathbb{E}[\|S^*Su\|])\|v\| \\ &\leq (\alpha\|u\| + \mathbb{E}[q\|Su\|])\|v\| \leq (\alpha\|u\| + \mathbb{E}\big[q^2\big]\|u\|)\|v\| \\ &\leq (\alpha + \mathbb{E}\big[q^2\big])\|u\|\|v\| \leq c\|u\|\|v\|. \end{aligned}$$

We use $\alpha > 0$ to show that a is coercive

$$\begin{aligned} a(u,u) &= \alpha\|u\|^2 + (\mathbb{E}[S^*S]u, u) = \alpha\|u\|^2 + \mathbb{E}[(S^*Su, u)] \\ &= \alpha\|u\|^2 + \mathbb{E}[(Su, Su)] = \alpha\|u\|^2 + \mathbb{E}\big[\|Su\|^2\big] \geq \alpha\|u\|^2. \end{aligned}$$

The boundedness of the right–hand side is straightforward to verify. Hence there exists a unique solution $u \in U$ of (7.3), (7.5) holds and the use of Assumption 7.2 shows

$$\|u\| = \|B^{-1}f\| \leq \frac{1}{\alpha}\|\mathbb{E}[S^*]y_d\| \leq \frac{1}{\alpha}\mathbb{E}[\|S^*y_d\|] \leq \frac{\mathbb{E}[q]}{\alpha}\|y_d\| \leq c\|y_d\|.$$

B is also clearly self–adjoint. A formal derivation for the existence and uniqueness of a minimizer of (7.1) follows from [73, Theorem 1.43]. The optimality conditions (7.3) follow from [73, Theorem 1.46] showing that J in (7.1) is Gateaux differentiable ([73, Definition 1.29]), that is directionally differentiable in all directions and the derivative is bounded. We have already shown the directionally differentiability and the boundedness is straightforward, since for all $v \in U$

$$|\partial_v J(u)| = |(Bu - f, v)| \leq c(\|u\| + 1)\|v\|. \qquad \square$$

For the rest of this chapter we simplify the notation by denoting the solution of (7.3) with u if not mentioned otherwise.

Remark 7.4. The optimality conditions are often written down with the help of the adjoint equation, see [73, Section 1.6]. The equivalent formulation of (7.3) is

$$\begin{aligned} y(\omega) &= S(\omega)u && \mathbb{P}\text{–almost surely}, && (7.6)\\ p(\omega) &= S^*(\omega)(y(\omega) - y_d) && \mathbb{P}\text{–almost surely}, && (7.7)\\ \alpha u + \mathbb{E}[p] &= 0. && && (7.8) \end{aligned}$$

Here (7.6) is called the state equation, (7.7) the adjoint equation and p the adjoint. The last equation (7.8) expresses that the gradient of J to be equal to zero. This formulation is often used in practice since (7.6)–(7.8) offers a convenient way to compute the gradient. $\diamond$

Spatial discretization. The discretization of (7.3) requires the approximation of the mean $\mathbb{E}$, the solution operator S and the adjoint operator S^*. We first write down a well–known perturbation result for errors in the operator or right–hand side.

Lemma 7.5 (Perturbation error)**.** Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ be Banach spaces. Let $A : \mathcal{X} \to \mathcal{B}$ be a linear, bounded, invertible operator whose inverse is bounded. Assume that $x \in \mathcal{X}$ and $b \in \mathcal{B}$ are such that

$$Ax = b.$$

Furthermore, assume that we have perturbations $\Delta A : \mathcal{X} \to \mathcal{B}$ linear, $\Delta x \in \mathcal{X}$ and $\Delta b \in \mathcal{B}$ such that

$$(A + \Delta A)(x + \Delta x) = b + \Delta b.$$

Let $A + \Delta A : \mathcal{X} \to \mathcal{B}$ be a linear, bounded, invertible operator whose inverse is bounded. Then the perturbation error in the solution is bounded as follows

$$\|\Delta x\|_{\mathcal{X}} \leq \|(A + \Delta A)^{-1}\|_{\mathrm{op}}(\|\Delta A x\|_{\mathcal{B}} + \|\Delta b\|_{\mathcal{B}}), \tag{7.9}$$

where $\|\cdot\|_{\mathrm{op}}$ denotes the operator norm.

Proof. We rewrite the perturbed system and use $Ax = b$

$$(A + \Delta A)\Delta x = -(A + \Delta A)x + b + \Delta b = (-Ax + b) - \Delta A x + \Delta b = -\Delta A x + \Delta b.$$

Now use that the inverse of $A + \Delta A$ exists and is bounded. □

We spatially discretize S and its adjoint S^*. We assume that these operators are pathwise bounded and their approximation error decays geometrically.

Assumption 7.6 (S_ℓ is regular)**.** Let $\ell \in \mathbb{N}$. The operator S_ℓ and its adjoint S_ℓ^* are $\mathbb{P}$–almost surely bounded

$$\begin{aligned} \|S_\ell(\omega)u\| &\leq q(\omega)\|u\| && \text{for all } u \in U, \\ \|S_\ell^*(\omega)y\| &\leq q(\omega)\|y\| && \text{for all } y \in Y. \end{aligned}$$

Furthermore, there exists the rate $\gamma > 0$ such that the following error estimates are $\mathbb{P}$–almost surely true

$$\begin{aligned} \|(S(\omega) - S_\ell(\omega))u\| &\leq q(\omega)2^{-\gamma\ell}\|u\| && \text{for all } u \in U, \\ \|(S^*(\omega) - S_\ell^*(\omega))y\| &\leq q(\omega)2^{-\gamma\ell}\|y\| && \text{for all } y \in Y. \end{aligned}$$

The scalar valued random variable q has bounded fourth moment. S_ℓ and S_ℓ^* are random variables. ◇

The previous assumption is often satisfied for an elliptic PDE with uniform mesh refinement, since a version of Theorem 2.39 and Theorem 2.40 holds. We write down the discretized version of the optimality conditions (7.3) such that for all $\ell \in \{1, \dots, L\}$

$$(\alpha I + \mathbb{E}[S_\ell^* S_\ell])u_\ell = \mathbb{E}[S_\ell^*]y_d,$$

where we again use abbreviations $Q_\ell := \mathbb{E}[S_\ell^* S_\ell]$ and $f_\ell := \mathbb{E}[S_\ell^*]y_d$

$$(\alpha I + Q_\ell)u_\ell = f_\ell.$$

We further abbreviate the operator $B_\ell := \alpha I + Q_\ell$

$$B_\ell u_\ell = f_\ell. \tag{7.10}$$

We derive the optimal control problem that has the optimality conditions (7.10).

Remark 7.7 (Spatially discretized optimal control problem)**.** We discretize the solution operator S to arive at the spatially discretized version of (7.1). For all $\ell \in \{1, \dots, L\}$ this problem reads

$$\begin{aligned} \min_{u \in U} \quad & J(u) := \frac{1}{2}\mathbb{E}\big[\|y_\ell(u) - y_d\|^2\big] + \frac{\alpha}{2}\|u\|^2 \\ \text{such that } \mathbb{P}\text{–a.s.} \quad & y_\ell(u,\omega) = S_\ell(\omega)u. \end{aligned} \tag{7.11}$$

In particular, we do not discretize the control u or the space U, which is the approach taken in [72]. We only replaced S with S_ℓ, thus we have to do the same for the optimality conditions in (7.3) to arrive at (7.10). ◇

We now verify that (7.10) and thus (7.11) have the unique solution u_ℓ. We further derive an error estimate.

Theorem 7.8 (Existence, uniqueness and error)**.** Let $\alpha > 0$, Assumption 7.2 and Assumption 7.6 be true. Then B_ℓ is bounded, self–adjoint and invertible with bounded inverse

$$\|{B_\ell}^{-1}\| \leq c. \tag{7.12}$$

Furthermore, for all $\ell \in \{1, \dots, L\}$ there exists a unique solution u_ℓ of (7.10) that satisfies

$$\|u_\ell\| \leq c\|y_d\|, \tag{7.13}$$

$$\|u_\ell - u\| \leq c2^{-\gamma\ell}\|y_d\|. \tag{7.14}$$

Proof. The existence of the inverse ${B_\ell}^{-1}$ and (7.12) follows from the Lax–Milgram Lemma similarly to the proof of Theorem 7.3. The bound (7.13) and also the self–adjointness follows similarly. For the error estimate (7.14) we use Lemma 7.5 with

$$\begin{aligned} A &= \alpha I + Q, & \Delta A &= Q_\ell - Q, \\ b &= f, & \Delta b &= f_\ell - f, \\ x &= u, & \Delta x &= u_\ell - u \end{aligned}$$

such that (7.9) reads

$$\begin{aligned} \|u_\ell - u\| &\leq \|(\alpha I + Q_\ell)^{-1}\|(\|(Q_\ell - Q)u\| + \|f_\ell - f\|) \\ &\leq c(\|(\mathbb{E}[S_\ell^* S_\ell] - \mathbb{E}[S^* S])u\| + \|(\mathbb{E}[S_\ell^*] - \mathbb{E}[S^*])y_d\|) \\ &= c[(I) + (II)]. \end{aligned}$$

Here we used that the inverse of $\alpha I + Q_\ell = B_\ell$ is bounded. We simplify (I) with the help of Assumption 7.6 and (7.4) to bound $\|u\|$

$$\begin{aligned} (I) &\leq \mathbb{E}[\|(S_\ell^* S_\ell - S^* S)u\|] \leq \mathbb{E}[\|S_\ell^*(S_\ell - S)u\|] + \mathbb{E}[\|(S_\ell^* - S^*)Su\|] \\ &\leq \mathbb{E}[q\|(S_\ell - S)u\|] + \mathbb{E}\big[q2^{-\gamma\ell}\|Su\|\big] \leq 2\mathbb{E}\big[q^2\big]2^{-\gamma\ell}\|u\| \\ &\leq c2^{-\gamma\ell}\|y_d\|. \end{aligned}$$

Similarly, we bound (II) using Assumption 7.6

$$(II) \leq \mathbb{E}[\|(S_\ell^* - S^*)y_d\|] \leq \mathbb{E}\big[q2^{-\gamma\ell}\|y_d\|\big] \leq c2^{-\gamma\ell}\|y_d\|.$$

We thus conclude (7.14) and the theorem. □

Similarly to before, from now throughout the remainder of this chapter we denote the solution of (7.10) with u_ℓ.

Remark 7.9 (Constant c)**.** The constant c in (7.12), (7.13) and (7.14) contains the factor α^{-1} and thus requires $\alpha > 0$. We do not mention the dependence on α for norm and error estimates of u or approximations thereof in the future. This constant also contains moments of q. ◇

7.2 Monte Carlo discretization

We continue with the discretization of $\mathbb{E}$ in (7.10) with standard MC. The fully discrete optimality conditions for $\ell \in \{1, \dots, L\}$ read

$$\left(\alpha I + \frac{1}{m}\sum_{i=1}^{m} S_\ell^*(\omega_i) S_\ell(\omega_i)\right) \widehat{u}_\ell^{\mathrm{MC}} = \frac{1}{m}\sum_{i=1}^{m} S_\ell^*(\omega_i) y_d,$$

where we abbreviate the operator and right–hand side to arrive at

$$\left(\alpha I + \widehat{Q}_\ell\right) \widehat{u}_\ell^{\mathrm{MC}} = \widehat{f}_\ell.$$

We further abbreviate the left operator $\widehat{B}_\ell := \alpha I + \widehat{Q}_\ell$ such that

$$\widehat{B}_\ell \widehat{u}_\ell^{\mathrm{MC}} = \widehat{f}_\ell. \tag{7.15}$$

We remark that we use the same i.i.d. samples for the discretization of the operator $\widehat{B}_\ell$ and the right–hand side $\widehat{f}_\ell$. We write down the corresponding optimal control problem.

Remark 7.10 (MC optimal control problem)**.** The optimality conditions (7.15) correspond to the optimal control problem (7.1) where we discretize both the expectation $\mathbb{E}$ and the solution operator S with S_ℓ for $\ell \in \{1, \dots, L\}$

$$\begin{aligned} \min_{u \in U} \quad & J(u) := \frac{1}{2}\frac{1}{m}\sum_{i=1}^{m} \|y_\ell(u, \omega_i) - y_d\|^2 + \frac{\alpha}{2}\|u\|^2 \\ \text{such that} \quad & y_\ell(u, \omega_i) = S_\ell(\omega_i) u \qquad \text{for all } i \in \{1, \dots, m\}. \end{aligned} \tag{7.16}$$

This is again a quadratic optimization problem with m linear equality constraints. ◇

We derive existence and uniqueness of a solution of (7.15) similar to Theorem 7.3 and Theorem 7.8. This result is not surprising, since the MC discretization of the expectation in (7.16) is a norm on the product space Y^m and the m equality constraints can be grouped together to a single equation with a block diagonal solution operator. We conclude that (7.16) is a standard deterministic optimal control problem with quadratic cost function and linear constraint, hence the standard theory applies.

Theorem 7.11 (MC existence, uniqueness and error)**.** Let $\alpha > 0$ and Assumption 7.6 be true. Then for all $\ell \in \{1, \dots, L\}$ the operator $\widehat{B}_\ell$ is self–adjoint and invertible with bounded inverse

$$\|\widehat{B}_\ell^{-1}\| \leq \frac{1}{\alpha} \leq c. \tag{7.17}$$

Furthermore, for all $\ell \in \{1, \dots, L\}$ there exists a unique solution $\widehat{u}_\ell^{\mathrm{MC}}$ of (7.15) that has bounded second moment and satisfies the estimates

$$\mathbb{E}\big[\|\widehat{u}_\ell^{\mathrm{MC}}\|^2\big] \leq c\|y_d\|^2, \tag{7.18}$$

$$\mathbb{E}\big[\|\widehat{u}_\ell^{\mathrm{MC}} - u_\ell\|^2\big] \leq c\frac{1}{m}\|y_d\|^2. \tag{7.19}$$

Proof. The invertibility of $\widehat{B}_\ell$, (7.17) and that $\widehat{B}_\ell$ is self–adjoint follows similarly to the proof of Theorem 7.8. We use (7.17) and Assumption 7.6 to calculate

$$\begin{aligned}\|\widehat{u}_\ell^{\mathrm{MC}}\| \leq \|\widehat{B}_\ell^{-1}\|\|\widehat{f}_\ell\| &\leq c\left\|\frac{1}{m}\sum_{i=1}^m S_\ell^*(\omega_i)y_d\right\| \\ &\leq c\frac{1}{m}\sum_{i=1}^m \|S_\ell^*(\omega_i)y_d\| \leq c\frac{1}{m}\sum_{i=1}^m q(\omega_i)\|y_d\|.\end{aligned}$$

We square this expression, apply the expectation operator and the Cauchy–Schwarz inequality

$$\begin{aligned}\mathbb{E}\big[\|\widehat{u}_\ell^{\mathrm{MC}}\|^2\big] &\leq c\mathbb{E}\left[\left(\frac{1}{m}\sum_{i=1}^m q(\omega_i)\right)^2\right]\|y_d\|^2 = c\|y_d\|^2\frac{1}{m^2}\sum_{i,j=1}^m \mathbb{E}[q(\omega_i)q(\omega_j)] \\ &\leq c\|y_d\|^2\frac{1}{m^2}\sum_{i,j=1}^m \mathbb{E}\big[q(\omega_i)^2\big]^{1/2}\mathbb{E}\big[q(\omega_j)^2\big]^{1/2} = c\|y_d\|^2\mathbb{E}\big[q^2\big].\end{aligned}$$

The result (7.18) then follows from Assumption 7.6 using that q has a bounded second moment. We now analyse the error introduced from sampling. We use Lemma 7.5 with

$$\begin{aligned} A &= \alpha I + Q_\ell, & \Delta A &= \widehat{Q}_\ell - Q_\ell, \\ b &= f_\ell, & \Delta b &= \widehat{f}_\ell - f_\ell, \\ x &= u_\ell, & \Delta x &= \widehat{u}_\ell^{\mathrm{MC}} - u_\ell. \end{aligned}$$

We insert these values into (7.9), square it, bound the inverse of $\widehat{B}_\ell = \alpha I + \widehat{Q}_\ell$ and use Young's inequality

$$\begin{aligned}\|\widehat{u}_\ell^{\mathrm{MC}} - u_\ell\|^2 &\leq \|(\alpha I + \widehat{Q}_\ell)^{-1}\|^2\Big(\|(\widehat{Q}_\ell - Q_\ell)u_\ell\| + \|\widehat{f}_\ell - f_\ell\|\Big)^2 \\ &\leq c\Big(\|(\widehat{Q}_\ell - Q_\ell)u_\ell\|^2 + \|\widehat{f}_\ell - f_\ell\|^2\Big) \\ &= c[(I) + (II)].\end{aligned}$$

We take the expectation, use the definition of the variance (2.2), the MC variance estimate, Assumption 7.6, bound the fourth moment of q and use (7.13) to bound $\|u_\ell\|^2$

$$\begin{aligned}\mathbb{E}[(I)] &= \mathbb{E}\left[\|\frac{1}{m}\sum_{i=1}^m S_\ell^*(\omega_i)S_\ell(\omega_i)u_\ell - \mathbb{E}[S_\ell^*S_\ell]u_\ell\|^2\right] = \mathbb{V}\left[\frac{1}{m}\sum_{i=1}^m S_\ell^*(\omega_i)S_\ell(\omega_i)u_\ell\right] \\ &= \frac{\mathbb{V}[S_\ell^*S_\ell u_\ell]}{m} \leq \frac{\mathbb{E}[\|S_\ell^*S_\ell u_\ell\|^2]}{m} \leq \frac{\mathbb{E}[q^4]\|u_\ell\|^2}{m} \leq c\frac{1}{m}\|y_d\|^2.\end{aligned}$$

We use exactly the same steps to estimate (II) from which (7.19) follows

$$\begin{aligned}\mathbb{E}[(II)] &= \mathbb{E}\left[\|\frac{1}{m}\sum_{i=1}^{m} S_\ell^*(\omega_i)y_d - \mathbb{E}[S_\ell^*]y_d\|^2\right] = \mathbb{V}\left[\frac{1}{m}\sum_{i=1}^{m} S_\ell^*(\omega_i)y_d\right] \\ &= \frac{\mathbb{V}[S_\ell^* y_d]}{m} \leq \frac{\mathbb{E}[\|S_\ell^* y_d\|^2]}{m} \leq \frac{\mathbb{E}[q^2]\|y_d\|^2}{m} \leq c\frac{1}{m}\|y_d\|^2. \qquad \square\end{aligned}$$

Remark 7.12 (Bias of MC)**.** It is crucial to remark that $\widehat{u}_\ell^{\mathrm{MC}}$ is in general not an unbiased estimator for u_ℓ since

$$\begin{aligned}\mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] &= \mathbb{E}\left[\left(\alpha I + \widehat{Q}_\ell\right)^{-1}\widehat{f}_\ell\right] \neq \mathbb{E}\left[\left(\alpha I + \widehat{Q}_\ell\right)^{-1}\right]\mathbb{E}\left[\widehat{f}_\ell\right] \\ &\neq \left(\alpha I + \mathbb{E}\left[\widehat{Q}_\ell\right]\right)^{-1}\mathbb{E}\left[\widehat{f}_\ell\right] = (\alpha I + Q_\ell)^{-1} f_\ell = u_\ell.\end{aligned}$$

This holds, since $\widehat{Q}_\ell$ is correlated with $\widehat{f}_\ell$ as both use the same samples. Furthermore, we cannot interchange the mean and the inverse. However, (7.19) shows that the bias decreases sufficiently fast with order at least $m^{-1/2}$. In fact, the actual order is m^{-1}, which we now verify. For a fixed $v \in U$ we use $\mathbb{E}\left[\widehat{f}_\ell\right] = f_\ell$ to obtain

$$(\mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] - u_\ell, v) = \mathbb{E}\left[(\widehat{B}_\ell^{-1}\widehat{f}_\ell - {B_\ell}^{-1}\widehat{f}_\ell, v)\right] = \mathbb{E}\left[({B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)\widehat{B}_\ell^{-1}\widehat{f}_\ell, v)\right].$$

Now crucially, we insert the mean zero term ${B_\ell}^{-1}(\widehat{B}_\ell - B_\ell){B_\ell}^{-1} f_\ell$ and use that ${B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)$ is self–adjoint to conclude

$$\begin{aligned}(\mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] - u_\ell, v) &= \mathbb{E}\left[({B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)(\widehat{B}_\ell^{-1}\widehat{f}_\ell - {B_\ell}^{-1} f_\ell), v)\right] \\ &= \mathbb{E}\left[(\widehat{B}_\ell^{-1}\widehat{f}_\ell - {B_\ell}^{-1} f_\ell, {B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)v)\right] \\ &= \mathbb{E}\left[(\widehat{u}_\ell^{\mathrm{MC}} - u_\ell, {B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)v)\right].\end{aligned}$$

Finally, we use the Cauchy–Schwarz inequality, the error estimate (7.19), we bound the right expression with a standard MC variance estimate and bound ${B_\ell}^{-1}$ to arrive at

$$\begin{aligned}|\mathbb{E}\left[(\widehat{u}_\ell^{\mathrm{MC}} - u_\ell, {B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)v)\right]| &\leq \mathbb{E}\big[\|\widehat{u}_\ell^{\mathrm{MC}} - u_\ell\|^2\big]^{1/2}\mathbb{E}\left[\|{B_\ell}^{-1}(\widehat{B}_\ell - B_\ell)v\|^2\right]^{1/2} \\ &\leq c\frac{1}{m^{1/2}}\|y_d\|\frac{1}{m^{1/2}}\|v\| = c\frac{1}{m}\|y_d\|\|v\|.\end{aligned}$$

Now set $v = \mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] - u_\ell \in U$ to conclude

$$\|\mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] - u_\ell\|^2 \leq c\frac{1}{m}\|y_d\|\|\mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] - u_\ell\|,$$

which shows that the bias of the MC estimator $\widehat{u}_\ell^{\mathrm{MC}}$ is at least of order m^{-1}. $\diamond$

In practice we have to solve the system (7.15). For PDE constraints the operator S_ℓ and thus $\widehat{Q}_\ell$ corresponds to a matrix that is often dense. Therefore, it is not practical to compute the operator $\widehat{B}_\ell = \alpha I + \widehat{Q}_\ell$ or its inverse. However, some iterative methods only require the application of the matrix $\widehat{B}_\ell$ to a vector. In this case, the MC method requires

a state solve applying S_ℓ and an adjoint solve applying S_ℓ^* for every realization. The convergence rate of such iterative methods is often linear and we derive results specifically for the *conjugate gradient* (CG) algorithm. The CG–algorithm we use is standard and given in [103, Algorithm 5.2]. This is a Krylov subspace method originally from the authors of [71] and well known in mathematical literature, see [131, Chapter 4], [78, Chapter 2] and [98, Chapter 2]. We denote the CG–approximation of $\widehat{u}_\ell^{\mathrm{MC}}$ with $\widehat{u}_\ell^{\mathrm{MC,\,CG}}$ starting the iteration with a zero vector.

Lemma 7.13 (Number of CG–iterations)**.** Let $\alpha > 0$ and Assumption 7.6 be true. For $\ell \in \{1, \dots, L\}$ let $\widehat{u}_\ell^{\mathrm{MC,\,CG}}$ be the CG–approximation of $\widehat{u}_\ell^{\mathrm{MC}}$. Then the convergence is linear such that after n iterations the error satisfies

$$\|\widehat{u}_\ell^{\mathrm{MC,\,CG}} - \widehat{u}_\ell^{\mathrm{MC}}\| \leq c \left[\frac{\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i)}{\alpha} \right]^{1/2} 2^{-\gamma_{\mathrm{CG}} n} \|\widehat{u}_\ell^{\mathrm{MC}}\|, \tag{7.20}$$

where the random rate γ_{CG} is lower bounded

$$\gamma_{\mathrm{CG}} \geq -\log_2 \left(1 - \left(\frac{\alpha}{\alpha + \frac{1}{m}\sum_{i=1}^m q^2(\omega_i)} \right)^{1/2} \right) > 0. \tag{7.21}$$

Furthermore, for all $\varepsilon \in (0, 1/e]$ the expected number of steps $\mathbb{E}[n]$ to achieve the error $\|\widehat{u}_\ell^{\mathrm{MC,\,CG}} - \widehat{u}_\ell^{\mathrm{MC}}\| \leq \varepsilon$ is logarithmic in ε

$$\mathbb{E}[n] \leq c(|\log_2(\varepsilon)| + \mathbb{E}\big[\|\widehat{u}_\ell^{\mathrm{MC}}\|^2\big]^{1/2}) \leq c|\log_2(\varepsilon)|. \tag{7.22}$$

Proof. We use the definition of the A–norm for a symmetric positive definite matrix A, which satisfies the well–known inequalities for all $z \in \mathbb{R}^n$

$$\|z\|_A^2 := z^T A z, \quad \|z\|_I^2 \leq \|A^{-1}\| \|z\|_A^2, \quad \|z\|_A^2 \leq \|A\| \|z\|_I^2.$$

We further use the estimate [103, Equation 5.36], where we use the starting point to be zero. According to this equation, after n steps the CG–approximation x^{CG} of x, where x solves $Ax = b$, satisfies

$$\|x^{\mathrm{CG}} - x\|_A \leq 2 \left(\frac{\sqrt{\|A\|\|A^{-1}\|} - 1}{\sqrt{\|A\|\|A^{-1}\|} + 1} \right)^n \|x\|_A.$$

We use this and define the condition number $\kappa := \|\widehat{B}_\ell\| \|\widehat{B}_\ell^{-1}\|$ to obtain

$$\|\widehat{u}_\ell^{\mathrm{MC,\,CG}} - \widehat{u}_\ell^{\mathrm{MC}}\|_I^2 \leq \|\widehat{B}_\ell^{-1}\| \|\widehat{u}_\ell^{\mathrm{MC,\,CG}} - \widehat{u}_\ell^{\mathrm{MC}}\|_{\widehat{B}_\ell}^2 \leq 4 \|\widehat{B}_\ell^{-1}\| \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2n} \|\widehat{u}_\ell^{\mathrm{MC}}\|_{\widehat{B}_\ell}^2.$$

A further computation then shows

$$\begin{aligned} \|\widehat{B}_\ell^{-1}\| \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2n} \|\widehat{u}_\ell^{\mathrm{MC}}\|_{\widehat{B}_\ell}^2 &\leq \|\widehat{B}_\ell^{-1}\| \|\widehat{B}_\ell\| \left(1 - 2\frac{1}{\sqrt{\kappa} + 1} \right)^{2n} \|\widehat{u}_\ell^{\mathrm{MC}}\|_I^2 \\ &= \kappa \left(1 - 2\frac{1}{\sqrt{\kappa} + 1} \right)^{2n} \|\widehat{u}_\ell^{\mathrm{MC}}\|_I^2. \end{aligned} \tag{7.23}$$

We now upper bound the condition number using $\|\widehat{B}_\ell^{-1}\| \leq 1/\alpha$ and Assumption 7.6

$$\begin{aligned}\kappa &\leq \frac{1}{\alpha} \sup_{\|v\|=1} \|(\alpha I + \widehat{Q}_\ell)v\| \\ &\leq \frac{1}{\alpha}\left(\alpha + \sup_{\|v\|=1}\left\|\frac{1}{m}\sum_{i=1}^m S_\ell^*(\omega_i)S_\ell(\omega_i)v\right\|\right) \\ &\leq \frac{\alpha + \frac{1}{m}\sum_{i=1}^m q^2(\omega_i)}{\alpha}.\end{aligned}$$

We insert this into (7.23) to obtain (7.20), since the linear convergence rate satisfies

$$\begin{aligned}2^{-\gamma_{\mathrm{CG}}} &\leq 1 - 2\frac{1}{\sqrt{\kappa}+1} \\ &\leq 1 - 2\frac{\alpha^{1/2}}{(\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i))^{1/2} + \alpha^{1/2}} \\ &\leq 1 - 2\frac{\alpha^{1/2}}{(\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i))^{1/2} + (\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i))^{1/2}}.\end{aligned}$$

We take the negative logarithm to conclude (7.21). For a fixed $\varepsilon \in (0, 1/e]$ we achieve $\|\widehat{u}_\ell^{\mathrm{MC,\,CG}} - \widehat{u}_\ell^{\mathrm{MC}}\| \leq \varepsilon$ using (7.20) if we choose

$$n = c\frac{-\log_2(\varepsilon) + \log_2(\|\widehat{u}_\ell^{\mathrm{MC}}\|) + \log_2\left(\left[\frac{\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i)}{\alpha}\right]^{1/2}\right)}{\gamma_{\mathrm{CG}}}.$$

Here we implicitly assume that $\|\widehat{u}_\ell^{\mathrm{MC}}\| > 0$, since we otherwise have to iterate at most once $n = 1$ for realizations with $\|\widehat{u}_\ell^{\mathrm{MC}}\| = 0$. We use that γ_{CG} is positive and the elementary inequality $\log_2(x) \leq 2x$ for $x > 0$ to conclude

$$n \leq c\frac{-\log_2(\varepsilon) + \|\widehat{u}_\ell^{\mathrm{MC}}\| + \left[\frac{\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i)}{\alpha}\right]^{1/2}}{\gamma_{\mathrm{CG}}}.$$

Now take the mean and use the Cauchy–Schwarz inequality for the last two terms

$$\mathbb{E}[n] \leq c|\log_2(\varepsilon)|\mathbb{E}\left[\frac{1}{\gamma_{\mathrm{CG}}}\right] + \left(\mathbb{E}\left[\|\widehat{u}_\ell^{\mathrm{MC}}\|^2\right]^{1/2} + \mathbb{E}\left[\frac{\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i)}{\alpha}\right]^{1/2}\right)\mathbb{E}\left[\frac{1}{\gamma_{\mathrm{CG}}^2}\right]^{1/2}.$$

We use Assumption 7.6, $\mathbb{E}[1/\gamma_{\mathrm{CG}}] \leq \mathbb{E}[1/\gamma_{\mathrm{CG}}^2]^{1/2}$ and $\varepsilon \in (0, 1/e]$ to conclude

$$\mathbb{E}\left[\widehat{t}\right] \leq c\Big(|\log_2(\varepsilon)| + \mathbb{E}\left[\|\widehat{u}_\ell^{\mathrm{MC}}\|^2\right]^{1/2}\Big)\mathbb{E}\left[\frac{1}{\gamma_{\mathrm{CG}}^2}\right]^{1/2}. \tag{7.24}$$

Since the norm of $\widehat{u}_\ell^{\mathrm{MC}}$ is bounded due to (7.18) of Lemma 7.13, it remains to show that the right–most term in (7.24) is bounded. The elementary inequality

$$-\log(1-x) \geq x \quad \text{for all } x \in [0, 1) \tag{7.25}$$

holds, which can easily be checked since equality holds for $x = 0$ and the derivative of the left side is larger than the derivative of the right side

$$\frac{1}{1-x} \geq 1 \quad \text{for all } x \in [0,1).$$

Using that both sides of (7.25) are non–negative we obtain

$$\log(1-x)^{-2} \leq x^{-2} \quad \text{for all } x \in [0,1).$$

This inequality also holds for the logarithm with base 2 if we add an appropriate constant. We combine this with (7.21) and Assumption 7.6 to finally bound

$$\mathbb{E}\big[1/\gamma_{\mathrm{CG}}^2\big] = \mathbb{E}\left[\log_2\left(1 - \left(\frac{\alpha}{\alpha + \frac{1}{m}\sum_{i=1}^m q^2(\omega_i)}\right)^{1/2}\right)^{-2}\right] \leq c\mathbb{E}\left[\frac{\alpha + \frac{1}{m}\sum_{\ell=1}^m q^2(\omega_i)}{\alpha}\right] \leq c.$$

□

The previous theorem allows us to write down the complexity of the MC estimator obtained from the CG–method.

Theorem 7.14 (Complexity of MC)**.** Let $\alpha > 0$, Assumption 7.2 and Assumption 7.6 be true. Let the cost increase for both an application of S_ℓ and its adjoint S_ℓ^* be at most geometric

$$\mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*] \leq c2^{\gamma_{\mathrm{Cost}}\ell} \qquad \text{for all } \ell \in \{1,\dots,L\}. \tag{7.26}$$

Then for all $\varepsilon \in (0, 1/e]$ there exists L and m such that $\mathbb{E}\Big[\|\widehat{u}_L^{\mathrm{MC,\,CG}} - u\|^2\Big] \leq \varepsilon^2$ with expected costs bounded by

$$\mathbb{W}\Big[\widehat{u}_L^{\mathrm{MC,\,CG}}\Big] \leq c|\log(\varepsilon)|\varepsilon^{-2-\gamma_{\mathrm{Cost}}/\gamma}. \tag{7.27}$$

Proof. The proof is an adaptation of the MC complexity Theorem 3.11 which includes the iteration error. We split up the error into three terms with the triangle inequality and Cauchy–Schwarz inequality with $(1,1,1)^T \in \mathbb{R}^3$

$$\begin{aligned}
\mathbb{E}\Big[\|\widehat{u}_L^{\mathrm{MC,\,CG}} - u\|^2\Big] &\leq \mathbb{E}\Big[\Big(\|\widehat{u}_L^{\mathrm{MC,\,CG}} - \widehat{u}_L^{\mathrm{MC}}\| + \|\widehat{u}_L^{\mathrm{MC}} - u_L\| + \|u_L - u\|\Big)^2\Big] \\
&\leq \left(\sum_{i=1}^3 1^2\right)\Big(\mathbb{E}\Big[\|\widehat{u}_L^{\mathrm{MC,\,CG}} - \widehat{u}_L^{\mathrm{MC}}\|^2\Big] + \mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MC}} - u_L\|^2\big] + \|u_L - u\|^2\Big) \\
&\leq 3((I) + (II) + (III)).
\end{aligned}$$

We require that the iteration error (I), the variance (II) and the bias error (III) are all smaller than $\varepsilon^2/9$. For $(III) \leq c\varepsilon^2$ the use of the error estimate (7.14) shows that we have to choose

$$L \geq \frac{-\log_2(\varepsilon)}{\gamma} + c. \tag{7.28}$$

We have $(II) \leq c\varepsilon^2$ using the error estimate (7.19) with the number of samples

$$m \geq c\varepsilon^{-2}. \tag{7.29}$$

Finally (7.22) shows that the expected iteration count is logarithmic w.r.t. ε

$$\mathbb{E}[n] \leq c|\log_2(\varepsilon)|. \tag{7.30}$$

We multiply the number of iterations (7.30) with the number of samples (7.29) and the cost on level L in (7.28). We then use the geometric cost increase (7.26) to obtain (7.27)

$$\mathbb{W}\left[\widehat{u}_L^{\mathrm{MC,\,CG}}\right] = \mathbb{E}[nm(\mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*])] \leq c\varepsilon^{-2}2^{\gamma_{\mathrm{Cost}}L}|\log_2(\varepsilon)| \leq c\varepsilon^{-2-\gamma_{\mathrm{Cost}}/\gamma}|\log_2(\varepsilon)|. \quad \square$$

The additional logarithmic factor in (7.27) compared to the result in Theorem 3.11 is purely a consequence of requiring a logarithmic number of iterations for the CG–algorithm. If the system $\widehat{B}_\ell$ can explicitly be formed and (7.15) can be solved with costs independently of L, then the logarithmic term disappears. Importantly, the optimal cost rate with $\gamma_{\mathrm{Cost}} = 0$ leads to the optimal (up to logarithmic factors) complexity $|\log_2(\varepsilon)|\varepsilon^{-2}$, which is not achieved by MC.

Remark 7.15 (Removing the logarithmic factor)**.** We outline a method for removing the logarithmic factor $|\log_2(\varepsilon)|$ in the asymptotic complexity (7.27). The idea is to use fewer samples at the start of the CG–iteration and geometrically increase their number. For simplicity and to avoid complications, assume that $m = 4^N$ for some $N \in \mathbb{N}$ and every CG–iteration reduces the error by half. We then start with $n = 1$ and 4^n samples denoting the corresponding estimator with $\widehat{u}_L^{\mathrm{MC},1}$ and its CG–approximation with $\widehat{u}_L^{\mathrm{MC,CG},1}$. We apply four iterations of the CG–method starting at zero and assuming $\|\widehat{u}_L^{\mathrm{MC},1}\| \leq 4$ to conclude

$$\|\widehat{u}_L^{\mathrm{MC,CG},1} - \widehat{u}_L^{\mathrm{MC},1}\| \leq 2 \times 0.5^4\|\widehat{u}_L^{\mathrm{MC},1}\| \leq \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC},1}\| \leq \frac{1}{2}.$$

We continue with $n = 2$ and 4^n samples to similarly define $\widehat{u}_L^{\mathrm{MC},2}$ and the corresponding CG–approximation with $\widehat{u}_L^{\mathrm{MC,CG},2}$. We apply four iterations of the CG–method starting with $\widehat{u}_L^{\mathrm{MC,CG},1}$ instead of a zero vector to conclude

$$\begin{aligned}\|\widehat{u}_L^{\mathrm{MC,CG},2} - \widehat{u}_L^{\mathrm{MC},2}\| &\leq 2 \times 0.5^4\|\widehat{u}_L^{\mathrm{MC,CG},1} - \widehat{u}_L^{\mathrm{MC},2}\| \\ &\leq \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC,CG},1} - \widehat{u}_L^{\mathrm{MC},1}\| + \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC},1} - \widehat{u}_L^{\mathrm{MC},2}\| \\ &\leq \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC,CG},1} - \widehat{u}_L^{\mathrm{MC},1}\| + \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC},1} - u_L\| + \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC},2} - u_L\|.\end{aligned}$$

Repeating this procedure sequentially with $n \in \{1, \dots, N\}$ shows the recursion

$$\|\widehat{u}_L^{\mathrm{MC,CG},n} - \widehat{u}_L^{\mathrm{MC},n}\| \leq \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC,CG},n-1} - \widehat{u}_L^{\mathrm{MC},n-1}\| + \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC},n-1} - u_L\| + \frac{1}{8}\|\widehat{u}_L^{\mathrm{MC},n} - u_L\|. \tag{7.31}$$

Now we make the assumption of $\|\widehat{u}_L^{\mathrm{MC},n} - u_L\| \leq 2^{-n}$, which is justified since we expect a RMSE of order 2^{-n} given 4^n samples. We then use induction over n to show $\|\widehat{u}_L^{\mathrm{MC,CG},n} - \widehat{u}_L^{\mathrm{MC},n}\| \leq 2^{-n}$ since (7.31) directly shows

$$\|\widehat{u}_L^{\mathrm{MC,CG},n} - \widehat{u}_L^{\mathrm{MC},n}\| \leq \frac{1}{8}(2^{-n+1} + 2^{-n+1} + 2^{-n}) \leq 2^{-n}.$$

We use that $\widehat{u}_L^{\mathrm{MC,CG},N} = \widehat{u}_L^{\mathrm{MC,\,CG}}$ and $\widehat{u}_L^{\mathrm{MC},N} = \widehat{u}_L^{\mathrm{MC}}$, which shows that the iteration error is of the same order as the RMSE $\|\widehat{u}_L^{\mathrm{MC}} - u_L\| \leq 2^{-N}$ since

$$\|\widehat{u}_L^{\mathrm{MC,\,CG}} - \widehat{u}_L^{\mathrm{MC}}\| \leq 2^{-N}. \tag{7.32}$$

We assume that the cost of applying one solution and adjoint operator is equal to one. Then the overall evaluation costs to achieve (7.32) is bounded by

$$4\sum_{n=1}^{N}(\mathbb{W}[S_L]+\mathbb{W}[S_L^*])4^n \leq 4\times 4^{N+1} \leq c4^N.$$

This cost however, is proportional to a single CG–iteration using all 4^N samples of $\widehat{u}_L^{\mathrm{MC}}$. Therefore, without making this argument rigorous, we believe that the logarithmic factor $|\log_2(\varepsilon)|$ can be removed from the asymptotic complexity by modifying the CG–method. ◇

7.3 Multilevel Monte Carlo for the control

We highlight why applying standard MLMC for the risk neutral optimal control problem (7.1) may lead to an ill–posed problem.

Remark 7.16 (Naive MLMC is ill–posed)**.** We discretize the the expectation with the standard MLMC estimator

$$\begin{aligned}\min_{u\in U} \quad & J(u) := \frac{1}{2}\sum_{\ell=1}^{L}\frac{1}{m_\ell}\sum_{i=1}^{m_\ell}\big(\|y_\ell(u,\omega_{i,\ell})-y_d\|^2-\|y_{\ell-1}(u,\omega_{i,\ell})-y_d\|^2\big)+\frac{\alpha}{2}\|u\|^2\\ \text{such that} \quad & y_\ell(u,\omega_{i,\ell}) = S_\ell(\omega_{i,\ell})u \qquad \text{for all } i\in\{1,\dots,m_\ell\}, \ell\in\{1,\dots,L\}.\end{aligned} \tag{7.33}$$

Here we defined $y_0 := y_d$. The cost function does not ensure $J \geq 0$, although this is the case for the original problem (7.1), the problem where only the solution operator S is discretized (7.11) and for the MC discretized version (7.16). Even worse, since we subtract a convex function from a convex function, it is not guaranteed that the resulting function is convex. We obtain such an example if we assume that $L := 2$, $m_1 = m_2 := 1$, $S_1(\omega_{1,1}) := S_2(\omega_{1,2}) := 0$ and $S_1(\omega_{1,2}) := I$. We further assume that $y_d := 0$. The cost function is then

$$J(u) = \frac{1}{2}\big(\|y_d\|^2-\|u-y_d\|^2+\|y_d\|^2\big)+\frac{\alpha}{2}\|u\|^2 = \frac{\alpha-1}{2}\|u\|^2,$$

which is strictly concave if $\alpha < 1$ and in this case does not have a minimizer. Hence, the problem (7.33) is ill–posed. We write down the first order necessary optimality conditions of (7.33)

$$\begin{aligned}&\left(\alpha I+\sum_{\ell=1}^{L}\frac{1}{m_\ell}\sum_{i=1}^{m_\ell}\big[S_\ell^*(\omega_{i,\ell})S_\ell(\omega_{i,\ell})-S_{\ell-1}^*(\omega_{i,\ell})S_{\ell-1}(\omega_{i,\ell})\big]\right)\widehat{u}_L^{\mathrm{MLMC}}\\ &=\sum_{\ell=1}^{L}\frac{1}{m_\ell}\sum_{i=1}^{m_\ell}\big[S_\ell^*(\omega_{i,\ell})-S_{\ell-1}^*(\omega_{i,\ell})\big]y_d.\end{aligned}$$

Here the operator on the left may not be invertible since even in the finite dimensional case the difference of two positive definite matrices is not necessarily positive definite. For the example in this remark, we have to solve

$$(\alpha-1)\widehat{u}_2^{\mathrm{MLMC}} = 0,$$

which does not have a unique solution for $\alpha = 1$ and for $\alpha < 1$ we compute the maximizer instead of the minimizer. Therefore, the naive application of MLMC is difficult to analyse since with positive probability we obtain an ill–posed minimization problem. ◇

The basic idea of MLC is to apply the MLMC telescoping sum idea to the control. We rewrite u_L as follows

$$u_L = \sum_{\ell=1}^{L} u_\ell - u_{\ell-1} = \sum_{\ell=1}^{L} \delta_\ell, \tag{7.34}$$

where we define $u_0 := 0$. We use the system (7.10) for ℓ and $\ell - 1$ to show that the increments $\delta_\ell := u_\ell - u_{\ell-1}$ satisfy

$$\begin{aligned} B_\ell \delta_\ell &= (\alpha I + Q_\ell)u_\ell - (\alpha I + Q_\ell)u_{\ell-1} \\ &= f_\ell - (\alpha I + Q_{\ell-1})u_{\ell-1} + (Q_{\ell-1} - Q_\ell)u_{\ell-1} \\ &= f_\ell - f_{\ell-1} - (Q_\ell - Q_{\ell-1})u_{\ell-1}. \end{aligned}$$

We use the difference notation $\Delta f_\ell := f_\ell - f_{\ell-1}$ and $\Delta Q_\ell := Q_\ell - Q_{\ell-1}$ to arrive at

$$B_\ell \delta_\ell = \Delta f_\ell - \Delta Q_\ell u_{\ell-1}. \tag{7.35}$$

Here we tacitly defined $S_0 := 0$ and $S_0^* := 0$ and thus $f_0 := 0$ and $Q_0 := 0$. We approximate (7.35) using MC in a recursive fashion

$$\begin{aligned} \left(\alpha I + \frac{1}{m_\ell}\sum_{i=1}^{m_\ell} S_\ell^*(\omega_{i,\ell})S_\ell(\omega_{i,\ell})\right)\widehat{\delta}_\ell = \frac{1}{m_\ell}\sum_{i=1}^{m_\ell}(S_\ell^*(\omega_{i,\ell}) - S_{\ell-1}^*(\omega_{i,\ell}))y_d \\ -\left(\frac{1}{m_\ell}\sum_{i=1}^{m_\ell}(S_\ell^*(\omega_{i,\ell})S_\ell(\omega_{i,\ell}) - S_{\ell-1}^*(\omega_{i,\ell})S_{\ell-1}(\omega_{i,\ell}))\right)\widehat{u}_{\ell-1}^{\mathrm{MLC}}. \end{aligned}$$

We abbreviate this as follows

$$\widehat{B}_\ell \widehat{\delta}_\ell = \Delta \widehat{f}_\ell - \Delta \widehat{Q}_\ell \widehat{u}_{\ell-1}^{\mathrm{MLC}}. \tag{7.36}$$

The controls $\widehat{u}_\ell^{\mathrm{MLC}}$ are defined as MC approximation of (7.34) with the help of $\widehat{\delta}_\ell$

$$\widehat{u}_\ell^{\mathrm{MLC}} := \sum_{j=1}^{\ell} \widehat{\delta}_j. \tag{7.37}$$

Importantly, we assume that samples are independent across levels in the sense that $\Delta\widehat{Q}_\ell, \Delta\widehat{f}_\ell$ are independent of $\Delta\widehat{Q}_j, \Delta\widehat{f}_j$ if $\ell \neq j$. However, the increments $\widehat{\delta}_\ell$ are not independent w.r.t. each other, since they depend on $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$ and thus in (7.37) on $\widehat{\delta}_{\ell-1}$. This is a key difference compared to the standard MLMC estimator and makes the error analysis more complicated.
We outline the advantages of the MLC approach. First of all, the system (7.36) is always invertible and in fact, uses the same operator $\widehat{B}_\ell$ as MC in (7.15). Secondly, we have a variance reduction for both terms in the right–hand side of (7.36) since the differences are of order $2^{-\gamma\ell}$. For a sampling based method the error $\widehat{\delta}_\ell - \delta_\ell$ is then of order $2^{-\gamma\ell}/m_\ell$ for all $\ell \in \{1, \dots, L\}$ and thus we need fewer samples on fine grids. We write down the corresponding optimal control problem to further explain the main idea.

Remark 7.17 (MLC optimal control problem)**.** We rewrite the optimality conditions (7.35) to redefine the desired state

$$B_\ell \delta_\ell = \mathbb{E}\big[S_\ell^*\big[\big(I - (S_\ell^*)^{-1}S_{\ell-1}^*\big)y_d - (S_\ell^*)^{-1}\big(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1}\big)u_{\ell-1}\big]\big] = \mathbb{E}[S_\ell^* y_{d,\ell}], \tag{7.38}$$

where the updated desired state $y_{d,\ell}$ is defined accordingly

$$y_{d,\ell} := \big(I - (S_\ell^*)^{-1} S_{\ell-1}^*\big) y_d - (S_\ell^*)^{-1}\big(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1}\big) u_{\ell-1}.$$

This expression is well defined if $(S_\ell^*)^{-1} S_{\ell-1}^*$ is well defined, which may or may not be true. It is straightforward to verify that (7.38) are the necessary and sufficient optimality conditions of

$$\begin{aligned} \min_{\delta \in U} \quad & J(\delta) := \frac{1}{2}\mathbb{E}\big[\|y_\ell(\delta) - y_{d,\ell}\|^2\big] + \frac{\alpha}{2}\|\delta\|^2 \\ \text{such that } \mathbb{P}\text{–a.s.} \quad & y_\ell(\delta, \omega) = S_\ell(\omega)\delta, \end{aligned} \tag{7.39}$$

where the desired state is now random and depends on the optimal control on the previous levels. We thus have to solve L consecutive convex optimal control problems starting with the coarsest level. We now discretize the expectation in (7.39) with MC and replace $u_{\ell-1}$ with $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$, which leads to the following optimal control problem for the ℓ–th level

$$\begin{aligned} \min_{\widehat{\delta} \in U} \quad & J(\widehat{\delta}) := \frac{1}{2}\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} \|y_\ell(\widehat{\delta}, \omega_{i,\ell}) - \widehat{y}_{d,\ell}(\omega_{i,\ell})\|^2 + \frac{\alpha}{2}\|\widehat{\delta}\|^2 \\ \text{such that} \quad & y_\ell(\widehat{\delta}, \omega_{i,\ell}) = S_\ell(\omega_{i,\ell})\widehat{\delta} \qquad \text{for all } i \in \{1, \dots, m_\ell\}, \end{aligned} \tag{7.40}$$

where the desired state is defined as

$$\widehat{y}_{d,\ell}(\omega) := \big(I - S_\ell^*(\omega)^{-1} S_{\ell-1}^*(\omega)\big) y_d - S_\ell^*(\omega)^{-1}\big(S_\ell^*(\omega) S_\ell(\omega) - S_{\ell-1}^*(\omega) S_{\ell-1}(\omega)\big)\widehat{u}_{\ell-1}^{\mathrm{MLC}}.$$

As it turns out, the optimality conditions of (7.40) are exactly (7.36). Importantly, the optimal control problem (7.40) remains convex. ◇

We now outline why this approach leads to a variance reduction. Thus assume for a moment that the random bound $q \in L^\infty$ in Assumption 7.6 and that the same bounds hold for the inverse $(S_\ell^*)^{-1}$, which we also assume exists. Previous error estimates in this chapter often include the factor $\|y_d\|^2$ and the same bounds also holds for (7.39) if we replace $\|y_d\|^2$ with $\|y_{d,\ell}\|^2$ since $q \in L^\infty$. Then a computation using Young's inequality, Assumption 7.6 and (7.13) shows

$$\begin{aligned} \mathbb{E}\big[\|y_{d,\ell}\|^2\big] &\leq \mathbb{E}\Big[\big(\|(S_\ell^*)^{-1}\|\|\big(S_\ell^* - S_{\ell-1}^*\big) y_d - \big(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1}\big) u_{\ell-1}\|\big)^2\Big] \\ &\leq c\mathbb{E}\big[\|\big(S_\ell^* - S_{\ell-1}^*\big) y_d\|^2 + \|\big(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1}\big) u_{\ell-1}\|^2\big] \\ &\leq c 2^{-2\gamma\ell}\|y_d\|^2. \end{aligned}$$

We further replace $\widehat{u}_\ell^{\mathrm{MC}}$ with $\widehat{\delta}_\ell$, u_ℓ with δ_ℓ and assume that $\widehat{u}_{\ell-1}^{\mathrm{MLC}} = u_{\ell-1}$ to obtain an improved bound from (7.19)

$$\mathbb{E}\Big[\|\widehat{\delta}_\ell - \delta_\ell\|^2\Big] \leq c\frac{1}{m_\ell}\mathbb{E}\big[\|y_{d,\ell}\|^2\big] \leq c\frac{2^{-2\gamma\ell}}{m_\ell}\|y_d\|^2.$$

This is the same error rate as for the standard MLMC estimator in (3.58) if $\mathbb{V}[Z_\ell - Z_{\ell-1}] \leq c2^{-2\gamma\ell}$. We therefore expect that $\widehat{u}_L^{\mathrm{MLC}}$ has similar complexity than the standard MLMC estimator. Notice that these results only hold if an exact approximation of $u_{\ell-1}$ is available but the approximation error of this quantity is also reduced using the same idea on a coarser level. These results also require stronger assumptions on the solution operator, its adjoint and the inverse of it, which we will not require in the remaining section.
The error analysis for the MLC estimator is more challenging than for the MLMC estimator since the increments $\widehat{\delta}_\ell$ are not independent of each other. We first provide a technical lemma which estimates some terms required for the error analysis.

Lemma 7.18 (Error estimates)**.** Let $\alpha > 0$ and Assumption 7.6 be true. Then for all $\ell \in \{1, \dots, L\}$

$$\mathbb{E}\Big[\|\Delta\widehat{f}_\ell - \Delta f_\ell\|^2\Big] \le c\frac{2^{-2\gamma\ell}}{m_\ell}\|y_d\|^2, \tag{7.41}$$

$$\mathbb{E}\Big[\|\Delta\widehat{Q}_\ell(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2\Big] \le c2^{-2\gamma\ell}\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big], \tag{7.42}$$

$$\mathbb{E}\Big[\|[\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2\Big] \le c\frac{1}{m_\ell}\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big], \tag{7.43}$$

$$\mathbb{E}\Big[\|(\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}\|^2\Big] \le c\frac{2^{-2\gamma\ell}}{m_\ell}\|y_d\|^2, \tag{7.44}$$

$$\mathbb{E}\Big[\|(\widehat{Q}_\ell - Q_\ell)\delta_\ell\|^2\Big] \le c\frac{2^{-2\gamma\ell}}{m_\ell}\|y_d\|^2. \tag{7.45}$$

Proof. (7.41): We use a standard MC error estimate and Assumption 7.6

$$\mathbb{E}\Big[\|\Delta\widehat{f}_\ell - \Delta f_\ell\|^2\Big] = \frac{\mathbb{V}\big[(S_\ell^* - S_{\ell-1}^*)y_d\big]}{m_\ell} \le \frac{\mathbb{E}\big[\|(S_\ell^* - S_{\ell-1}^*)y_d\|^2\big]}{m_\ell} \le \frac{\mathbb{E}[q^2]2^{-2\gamma\ell}\|y_d\|^2}{m_\ell}.$$

(7.42): We use the triangle inequality and Assumption 7.6 to conclude that for all $v \in U$

$$\|(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1})v\| \le \|S_\ell^*(S_\ell - S_{\ell-1})v\| + \|(S_\ell^* - S_{\ell-1}^*)S_{\ell-1}v\| \le 2q^2 2^{-\gamma\ell}\|v\|. \tag{7.46}$$

We remark that $S_\ell^*(\omega_{i,\ell})S_\ell(\omega_{i,\ell}) - S_{\ell-1}^*(\omega_{i,\ell})S_{\ell-1}(\omega_{i,\ell})$ is independent of $\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}$ and thus we estimate

$$\begin{aligned}
\|\Delta\widehat{Q}_\ell(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\| &\le \frac{1}{m_\ell}\sum_{i=1}^{m_\ell}\|\big[S_\ell^*(\omega_{i,\ell})S_\ell(\omega_{i,\ell}) - S_{\ell-1}^*(\omega_{i,\ell})S_{\ell-1}(\omega_{i,\ell})\big]\big)(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\| \\
&\le c\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} q^2(\omega_{i,\ell})2^{-\gamma\ell}\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\| \\
&= c\left(\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} q^2(\omega_{i,\ell})\right)2^{-\gamma\ell}\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|.
\end{aligned}$$

We square this expression, take the expectation, use the independence of $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$ and q, the Cauchy–Schwarz inequality and Assumption 7.6

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{1}{m_\ell}\sum_{i=1}^{m_\ell} q^2(\omega_{i,\ell})\right)^2 2^{-2\gamma\ell}\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\right] &\le 2^{-2\gamma\ell}\mathbb{E}\big[q^4\big]\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big] \\
&\le c2^{-2\gamma\ell}\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big].
\end{aligned}$$

(7.43): This expression is the difference of two solutions of a linear system with the same right–hand side. We apply Lemma 7.5 with

$$\begin{aligned}
A &= \alpha I + Q_\ell, & \Delta A &= \widehat{Q}_\ell - Q_\ell, \\
b &= \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}, & \Delta b &= 0, \\
x &= B_\ell{}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}), & \Delta x &= [\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}).
\end{aligned}$$

The perturbation result (7.9) reads

$$\begin{aligned}\|[\widehat{B}_\ell^{-1} - {B_\ell}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\| &\leq \|\widehat{B}_\ell^{-1}\|\|(\widehat{Q}_\ell - Q_\ell){B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\| \\ &\leq c\|(\widehat{Q}_\ell - Q_\ell){B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|.\end{aligned}$$

Here we used the bound $\|\widehat{B}_\ell^{-1}\| \leq 1/\alpha \leq c$. We square this and take the conditional expectation noting that $\widehat{Q}_\ell$ is independent of $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$ to obtain

$$\mathbb{E}\Big[\|(\widehat{Q}_\ell - Q_\ell){B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2|\widehat{u}_{\ell-1}^{\mathrm{MLC}}\Big] = \mathbb{V}\Big[\widehat{Q}_\ell{B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})|\widehat{u}_{\ell-1}^{\mathrm{MLC}}\Big].$$

We use the standard MC variance estimate, Assumption 7.6 and the bound $\|{B_\ell}^{-1}\| \leq c$

$$\begin{aligned}\mathbb{V}\Big[\widehat{Q}_\ell{B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})|\widehat{u}_{\ell-1}^{\mathrm{MLC}}\Big] &= \frac{\mathbb{V}\big[S_\ell^* S_\ell {B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})|\widehat{u}_{\ell-1}^{\mathrm{MLC}}\big]}{m_\ell} \\ &\leq \frac{\mathbb{E}\big[\|S_\ell^* S_\ell {B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2|\widehat{u}_{\ell-1}^{\mathrm{MLC}}\big]}{m_\ell} \\ &\leq \frac{\mathbb{E}[q^4]\|{B_\ell}^{-1}(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2}{m_\ell} \\ &\leq c\frac{\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2}{m_\ell}.\end{aligned}$$

Now use the law of total expectation $\mathbb{E}[\cdot] = \mathbb{E}\big[\mathbb{E}\big[\cdot|\widehat{u}_{\ell-1}^{\mathrm{MLC}}\big]\big]$ to conclude (7.43).
(7.44): We use the standard MC variance estimate, (7.46) with $v = u_{\ell-1} \in U$ and Assumption 7.6

$$\begin{aligned}\mathbb{E}\Big[\|(\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}\|^2\Big] = \mathbb{V}\Big[(\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}\Big] &= \frac{\mathbb{V}\big[(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1})u_{\ell-1}\big]}{m_\ell} \\ &\leq \frac{\mathbb{E}\big[\|(S_\ell^* S_\ell - S_{\ell-1}^* S_{\ell-1})u_{\ell-1}\|^2\big]}{m_\ell} \\ &\leq c\frac{\mathbb{E}[q^4]2^{-2\gamma\ell}\|u_{\ell-1}\|^2}{m_\ell} \\ &\leq c\frac{2^{-2\gamma\ell}}{m_\ell}\|u_{\ell-1}\|^2.\end{aligned}$$

We then use (7.13) to bound $u_{\ell-1}$ in the previous equation

$$\|u_{\ell-1}\|^2 \leq c\|y_d\|^2.$$

(7.45): We use a MC variance estimate and Assumption 7.6

$$\mathbb{E}\Big[\|(\widehat{Q}_\ell - Q_\ell)\delta_\ell\|^2\Big] = \mathbb{V}\Big[\widehat{Q}_\ell\delta_\ell\Big] = \frac{\mathbb{V}[S_\ell^* S_\ell\delta_\ell]}{m_\ell} \leq \frac{\mathbb{E}[\|S_\ell^* S_\ell\delta_\ell\|^2]}{m_\ell} \leq \frac{\mathbb{E}[q^4]}{m_\ell}\|\delta_\ell\|^2.$$

The result then follows since q has bounded fourth moment and the error estimate (7.14) shows

$$\|\delta_\ell\| = \|u_\ell - u_{\ell-1}\| \leq \|u_\ell - u\| + \|u_{\ell-1} - u\| \leq c2^{-\gamma\ell}\|y_d\|. \qquad \square$$

The next step is again a technical lemma that derives an expression for a mixed term. This estimate is needed later to ensure that the MSE decays fast enough.

Lemma 7.19 (Mixed estimate)**.** Let $\alpha > 0$ and Assumption 7.6 be true. Then for all $\ell \in \{1, \dots, L\}$ the estimate holds

$$|\mathbb{E}\Big[(\widehat{\delta}_\ell - \delta_\ell, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big]| \leq c\frac{2^{-2\gamma\ell}}{m_\ell}\|y_d\|^2 + c\bigg(\frac{1}{m_\ell} + 2^{-\gamma\ell}\bigg)\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]. \quad (7.47)$$

Proof. We use the definition of $\widehat{\delta}_\ell$ and δ_ℓ in (7.35) and (7.36) to obtain

$$\begin{aligned}
\widehat{\delta}_\ell - \delta_\ell &= \widehat{B}_\ell^{-1}(\Delta\widehat{f}_\ell - \widehat{Q}_\ell\widehat{u}_{\ell-1}^{\mathrm{MLC}}) - {B_\ell}^{-1}(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) \\
&\quad - \widehat{B}_\ell^{-1}(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) + \widehat{B}_\ell^{-1}(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) \\
&= \widehat{B}_\ell^{-1}(\Delta\widehat{f}_\ell - \Delta f_\ell) - \widehat{B}_\ell^{-1}(\Delta\widehat{Q}_\ell\widehat{u}_{\ell-1}^{\mathrm{MLC}} - \Delta Q_\ell u_{\ell-1}) \\
&\quad - ({B_\ell}^{-1} - \widehat{B}_\ell^{-1})(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) \\
&= \widehat{B}_\ell^{-1}(\Delta\widehat{f}_\ell - \Delta f_\ell) - \widehat{B}_\ell^{-1}(\Delta\widehat{Q}_\ell\widehat{u}_{\ell-1}^{\mathrm{MLC}} - \Delta\widehat{Q}_\ell u_{\ell-1}) \\
&\quad - \widehat{B}_\ell^{-1}(\Delta\widehat{Q}_\ell u_{\ell-1} - \Delta Q_\ell u_{\ell-1}) - ({B_\ell}^{-1} - \widehat{B}_\ell^{-1})(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) \\
&= (I) + (II) + (III) + (IV).
\end{aligned} \quad (7.48)$$

Let us now estimate the term (I). We use that $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$ is independent of $\widehat{f}_\ell$ to insert the mean zero expression ${B_\ell}^{-1}(\Delta\widehat{f}_\ell - \Delta f_\ell)$ into the scalar product

$$\mathbb{E}\Big[(\widehat{B}_\ell^{-1}(\Delta\widehat{f}_\ell - \Delta f_\ell), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big] = \mathbb{E}\Big[([\widehat{B}_\ell^{-1} - {B_\ell}^{-1}](\Delta\widehat{f}_\ell - \Delta f_\ell), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big].$$

We use that adjoints satisfy $(A+B)^* = A^* + B^*$, $(A^{-1})^* = (A^*)^{-1}$ and that both $\widehat{B}_\ell^{-1}$ and ${B_\ell}^{-1}$ are self–adjoint. We combine this with the Cauchy–Schwarz inequality

$$\begin{aligned}
&|\mathbb{E}\Big[([\widehat{B}_\ell^{-1} - {B_\ell}^{-1}](\Delta\widehat{f}_\ell - \Delta f_\ell), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big]| \\
&= |\mathbb{E}\Big[(\Delta\widehat{f}_\ell - \Delta f_\ell, [\widehat{B}_\ell^{-1} - {B_\ell}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}))\Big]| \\
&\leq \mathbb{E}\Big[\|\Delta\widehat{f}_\ell - \Delta f_\ell\|^2\Big]^{1/2}\mathbb{E}\Big[\|[\widehat{B}_\ell^{-1} - {B_\ell}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2\Big]^{1/2}.
\end{aligned}$$

We now use (7.41) and (7.43) to conclude

$$|\mathbb{E}\big[((I), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| \leq c\frac{2^{-\gamma\ell}}{m_\ell}\|y_d\|\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]^{1/2}. \quad (7.49)$$

Let us now estimate (II). We use the bound $\|\widehat{B}_\ell^{-1}\| \leq c$ and the Cauchy–Schwarz inequality

$$\begin{aligned}
|\mathbb{E}\big[((II), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| &= |\mathbb{E}\Big[(\widehat{B}_\ell^{-1}\Delta\widehat{Q}_\ell(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big]| \\
&\leq c\mathbb{E}\Big[\|\Delta\widehat{Q}_\ell(\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2\Big]^{1/2}\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]^{1/2}.
\end{aligned}$$

We combine this with (7.42) to conclude

$$|\mathbb{E}\big[((II), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| \leq c2^{-\gamma\ell}\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]. \quad (7.50)$$

Let us continue with the term (III). We use that ${B_\ell}^{-1}(\Delta\widehat{Q}_\ell u_{\ell-1} - \Delta Q_\ell u_{\ell-1})$ has mean zero and is independent of $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$ to conclude

$$\begin{aligned}
\mathbb{E}\big[((III), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big] &= \mathbb{E}\Big[(-\widehat{B}_\ell^{-1}(\Delta\widehat{Q}_\ell u_{\ell-1} - \Delta Q_\ell u_{\ell-1}), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big] \\
&= \mathbb{E}\Big[([{B_\ell}^{-1} - \widehat{B}_\ell^{-1}](\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big].
\end{aligned}$$

Similarly to the calculation of (I), we use the properties of the adjoints and inverses. Then we apply the Cauchy–Schwarz inequality

$$\begin{aligned}
&|\mathbb{E}\Big[([B_\ell{}^{-1} - \widehat{B}_\ell^{-1}](\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big]| \\
&= |\mathbb{E}\Big[((\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}, [B_\ell{}^{-1} - \widehat{B}_\ell^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}))\Big]| \\
&\leq \mathbb{E}\Big[\|(\Delta\widehat{Q}_\ell - \Delta Q_\ell)u_{\ell-1}\|^2\Big]^{1/2}\mathbb{E}\Big[\|[B_\ell{}^{-1} - \widehat{B}_\ell^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2\Big]^{1/2}.
\end{aligned}$$

We use (7.43) and (7.44) to bound this term

$$\begin{aligned}
|\mathbb{E}\big[((III), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| &\leq c\frac{2^{-\gamma\ell}}{(m_\ell)^{1/2}}\|y_d\|\frac{\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]^{1/2}}{(m_\ell)^{1/2}} \\
&= c\frac{2^{-\gamma\ell}}{m_\ell}\|y_d\|\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]^{1/2}.
\end{aligned} \tag{7.51}$$

We now derive an expression for (IV). A calculation and $B_\ell{}^{-1}(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) = \delta_\ell$ from (7.35) shows

$$\begin{aligned}
(IV) &= [\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) = \widehat{B}_\ell^{-1}[B_\ell - \widehat{B}_\ell]B_\ell{}^{-1}(\Delta f_\ell - \Delta Q_\ell u_{\ell-1}) \\
&= \widehat{B}_\ell^{-1}(Q_\ell - \widehat{Q}_\ell)\delta_\ell.
\end{aligned}$$

We use that $B_\ell{}^{-1}(Q_\ell - \widehat{Q}_\ell)\delta_\ell$ has mean zero and is independent of $\widehat{u}_{\ell-1}^{\mathrm{MLC}}$, which allows us to subtract it

$$\begin{aligned}
\mathbb{E}\big[((IV), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big] &= \mathbb{E}\Big[(\widehat{B}_\ell^{-1}(Q_\ell - \widehat{Q}_\ell)\delta_\ell, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big] \\
&= \mathbb{E}\Big[([\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](Q_\ell - \widehat{Q}_\ell)\delta_\ell, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big].
\end{aligned}$$

Similarly to before, use the properties of adjoints and the Cauchy–Schwarz inequality

$$\begin{aligned}
&|\mathbb{E}\Big[([\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](Q_\ell - \widehat{Q}_\ell)\delta_\ell, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big]| \\
&= |\mathbb{E}\Big[((Q_\ell - \widehat{Q}_\ell)\delta_\ell, [\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}))\Big]| \\
&\leq \mathbb{E}\Big[\|(Q_\ell - \widehat{Q}_\ell)\delta_\ell\|^2\Big]^{1/2}\mathbb{E}\Big[\|[\widehat{B}_\ell^{-1} - B_\ell{}^{-1}](\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\|^2\Big]^{1/2}.
\end{aligned}$$

We use (7.43) and (7.45) to bound this

$$|\mathbb{E}\big[((IV), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| \leq c\frac{2^{-\gamma\ell}}{m_\ell}\|y_d\|\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]^{1/2}. \tag{7.52}$$

We combine (7.48) with (7.49), (7.50), (7.51) and (7.52) to obtain

$$\begin{aligned}
|\mathbb{E}\Big[(\widehat{\delta}_\ell - \delta_\ell, \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\Big]| &\leq |\mathbb{E}\big[((I), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| + |\mathbb{E}\big[((II), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| \\
&\quad + |\mathbb{E}\big[((III), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| + |\mathbb{E}\big[((IV), \widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1})\big]| \\
&\leq c\frac{2^{-\gamma\ell}}{m_\ell}\|y_d\|\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big]^{1/2} + c2^{-\gamma\ell}\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}} - u_{\ell-1}\|^2\big].
\end{aligned}$$

The result (7.47) now follows from Young's inequality

$$\frac{2^{-\gamma\ell}}{m_\ell}\|y_d\|\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1}\|^2\big]^{1/2} = \|y_d\|\frac{2^{-\gamma\ell}}{(m_\ell)^{1/2}}\left(\frac{\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1}\|^2\big]}{m_\ell}\right)^{1/2}$$
$$\leq c\|y_d\|^2\frac{2^{-2\gamma\ell}}{m_\ell} + c\frac{\mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1}\|^2\big]}{m_\ell}. \qquad \square$$

Before we continue, we verify that the increment $\widehat{\delta}_\ell$ has bounded second moment and satisfies an error estimate.

Lemma 7.20 (Bound and error estimate for $\widehat{\delta}_\ell$)**.** Let $\alpha > 0$ and Assumption 7.6 be true. Then for all $\ell \in \{1,\dots,L\}$ the increments $\widehat{\delta}_\ell$ have bounded second moment and satisfy the error estimate

$$\mathbb{E}\Big[\|\widehat{\delta}_\ell\|^2\Big] \leq c2^{-2\gamma\ell}(\|y_d\|^2 + \mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1}\|^2\big]), \tag{7.53}$$
$$\mathbb{E}\Big[\|\widehat{\delta}_\ell-\delta_\ell\|^2\Big] \leq c2^{-2\gamma\ell}\left(\frac{1}{m_\ell}\|y_d\|^2 + \mathbb{E}\big[\|\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1}\|^2\big]\right). \tag{7.54}$$

Proof. We first show (7.54). We use the perturbation result (7.9) with

$$A = \alpha I + Q_\ell, \qquad \Delta A = \widehat{Q}_\ell - Q_\ell,$$
$$b = \Delta f_\ell - \Delta Q_\ell u_{\ell-1}, \qquad \Delta b = \Delta\widehat{f}_\ell - \Delta\widehat{Q}_\ell\widehat{u}_{\ell-1}^{\mathrm{MLC}} - \Delta f_\ell + \Delta Q_\ell u_{\ell-1},$$
$$x = \delta_\ell, \qquad \Delta x = \widehat{\delta}_\ell - \delta_\ell.$$

We combine this with the triangle inequality and the bound $\|\widehat{B}_\ell^{-1}\| \leq c$

$$\|\widehat{\delta}_\ell-\delta_\ell\| \leq \|\widehat{B}_\ell^{-1}\|(\|(\widehat{Q}_\ell-Q_\ell)\delta_\ell\| + \|\Delta\widehat{f}_\ell-\Delta f_\ell\| + \|\Delta\widehat{Q}_\ell\widehat{u}_{\ell-1}^{\mathrm{MLC}} - \Delta Q_\ell u_{\ell-1}\|)$$
$$\leq c(\|(\widehat{Q}_\ell-Q_\ell)\delta_\ell\| + \|\Delta\widehat{f}_\ell-\Delta f_\ell\| + \|\Delta\widehat{Q}_\ell(\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1})\| + \|(\Delta\widehat{Q}_\ell-\Delta Q_\ell)u_{\ell-1}\|).$$

We square this, apply the Cauchy–Schwarz inequality with $(1,1,1,1)^T \in \mathbb{R}^4$ and take the expectation to conclude

$$\mathbb{E}\Big[\|\widehat{\delta}_\ell-\delta_\ell\|^2\Big] \leq 4c(\mathbb{E}\Big[\|(\widehat{Q}_\ell-Q_\ell)\delta_\ell\|^2\Big] + \mathbb{E}\Big[\|\Delta\widehat{f}_\ell-\Delta f_\ell\|^2\Big]$$
$$+ \mathbb{E}\Big[\|\Delta\widehat{Q}_\ell(\widehat{u}_{\ell-1}^{\mathrm{MLC}}-u_{\ell-1})\|^2\Big] + \mathbb{E}\Big[\|(\Delta\widehat{Q}_\ell-\Delta Q_\ell)u_{\ell-1}\|^2\Big]).$$

These expressions are estimated in Lemma 7.18 from which we directly deduce (7.54). The bound (7.53) follows using the triangle inequality and the Cauchy–Schwarz inequality

$$\mathbb{E}\Big[\|\widehat{\delta}_\ell\|^2\Big] \leq c\mathbb{E}\Big[\|\widehat{\delta}_\ell-\delta_\ell\|^2\Big] + c\mathbb{E}\big[\|\delta_\ell\|^2\big] = c\mathbb{E}\Big[\|\widehat{\delta}_\ell-\delta_\ell\|^2\Big] + c\|u_\ell-u\|^2 + c\|u_{\ell-1}-u\|^2.$$

Now use the result (7.54), $m_\ell \geq 1$ and the error estimate (7.14) to obtain (7.53). $\square$

Lemma 7.20 allows us to derive an error estimate for the MLC estimator $\widehat{u}_\ell^{\mathrm{MLC}}$ as well as a bound for the second moment of it.

Theorem 7.21 (Bound and error estimate for $\widehat{u}_\ell^{\mathrm{MLC}}$)**.** Let $\alpha > 0$ and Assumption 7.6 be true. Then for all $L \in \mathbb{N}$ the bound and error estimate hold

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}}\|^2\big] \leq c\left(1 + \sum_{\ell=1}^{L} \frac{2^{-2\gamma\ell}}{m_\ell} \prod_{j=\ell+1}^{L} \left[1 + c\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right]\right) \|y_d\|^2, \quad (7.55)$$

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] \leq c\left(\sum_{\ell=1}^{L} \frac{2^{-2\gamma\ell}}{m_\ell} \prod_{j=\ell+1}^{L} \left[1 + c\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right]\right) \|y_d\|^2. \quad (7.56)$$

In particular, if the number of samples satisfy

$$\sum_{\ell=1}^{L} \frac{1}{m_\ell} < c \quad (7.57)$$

with a constant c independent of L, then it holds

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}}\|^2\big] \leq c\|y_d\|^2, \quad (7.58)$$

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] \leq c\left(\sum_{\ell=1}^{L} \frac{2^{-2\gamma\ell}}{m_\ell}\right) \|y_d\|^2. \quad (7.59)$$

Proof. We use the recursive definitions $u_L = \delta_L + u_{L-1}$ and $\widehat{u}_L^{\mathrm{MLC}} = \widehat{\delta}_L + \widehat{u}_{L-1}^{\mathrm{MLC}}$ together with the properties of the scalar product to obtain

$$\begin{aligned}\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] &= \mathbb{E}\Big[\|\widehat{\delta}_L - \delta_L\|^2\Big] + 2\mathbb{E}\Big[(\widehat{\delta}_L - \delta_L, \widehat{u}_{L-1}^{\mathrm{MLC}} - u_{L-1})\Big] + \mathbb{E}\big[\|\widehat{u}_{L-1}^{\mathrm{MLC}} - u_{L-1}\|^2\big] \\ &\leq \mathbb{E}\Big[\|\widehat{\delta}_L - \delta_L\|^2\Big] + 2|\mathbb{E}\Big[(\widehat{\delta}_L - \delta_L, \widehat{u}_{L-1}^{\mathrm{MLC}} - u_{L-1})\Big]| + \mathbb{E}\big[\|\widehat{u}_{L-1}^{\mathrm{MLC}} - u_{L-1}\|^2\big].\end{aligned}$$

We now use Lemma 7.19 and Lemma 7.20 to obtain the recursion

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] \leq c_0 \frac{2^{-2\gamma L}}{m_L} \|y_d\|^2 + \left[1 + c_0\left(\frac{1}{m_L} + 2^{-\gamma L} + 2^{-2\gamma L}\right)\right] \mathbb{E}\big[\|\widehat{u}_{L-1}^{\mathrm{MLC}} - u_{L-1}\|^2\big]. \quad (7.60)$$

We explicitly use the constant c_0 instead of c to verify that this constant does not depend on L. We apply induction over L to verify

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] \leq c_0\left(\sum_{\ell=1}^{L} \frac{2^{-2\gamma\ell}}{m_\ell} \prod_{j=\ell+1}^{L} \left[1 + c_0\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right]\right) \|y_d\|^2. \quad (7.61)$$

Notice that $\mathbb{E}\big[\|\widehat{u}_0^{\mathrm{MLC}} - u_0\|^2\big] = 0$ by definition and thus the recursion (7.60) for $L = 1$ is valid. We use the induction hypothesis assuming that (7.61) is valid for $L - 1$

$$\begin{aligned}&\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] \leq c_0 \frac{2^{-2\gamma L}}{m_L} \|y_d\|^2 \\ &+ \left[1 + c_0\left(\frac{1}{m_L} + 2^{-\gamma L} + 2^{-2\gamma L}\right)\right] c_0\left(\sum_{\ell=1}^{L-1} \frac{2^{-2\gamma\ell}}{m_\ell} \prod_{j=\ell+1}^{L-1} \left[1 + c_0\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right]\right) \|y_d\|^2 \\ &= c_0 \frac{2^{-2\gamma L}}{m_L} \|y_d\|^2 + c_0\left(\sum_{\ell=1}^{L-1} \frac{2^{-2\gamma\ell}}{m_\ell} \prod_{j=\ell+1}^{L} \left[1 + c_0\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right]\right) \|y_d\|^2 \\ &= c_0\left(\sum_{\ell=1}^{L} \frac{2^{-2\gamma\ell}}{m_\ell} \prod_{j=\ell+1}^{L} \left[1 + c_0\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right]\right) \|y_d\|^2.\end{aligned}$$

Here we used that the empty product is equal to one for the last equation. This shows the result (7.56) for all L. The infinite product converges to a finite value if the corresponding sequence is summable, that is for a sequence of summable, non–negative numbers $(a_n)_{n=1}^{\infty}$ we have

$$\begin{aligned}\log\left(\prod_{n=1}^{\infty}(1+a_n)\right) &= \log\left(\lim_{N\to+\infty}\prod_{n=1}^{N}(1+a_n)\right) = \lim_{N\to+\infty}\sum_{n=1}^{N}\log(1+a_n)\\ &= \sum_{n=1}^{\infty}\log(1+a_n) \leq \sum_{n=1}^{\infty} a_n < +\infty.\end{aligned}$$

We thus have to ensure that

$$\sum_{\ell=1}^{L}\left(\frac{1}{m_\ell} + 2^{-\gamma\ell} + 2^{-2\gamma\ell}\right) \leq c$$

independently of L. This is satisfied if the additional assumption (7.57) is true and using $\gamma > 0$ from Assumption 7.6. This shows (7.59). The remaining bounds (7.55) and (7.58) can be proved similarly to (7.53) in Lemma 7.20 by bounding u_L using (7.13). □

We remark that (7.59) is equal to the bound of the variance of the MLMC estimator in (3.58) by estimating $\mathbb{V}[Z_\ell - Z_{\ell-1}] \leq c2^{-2\gamma\ell}\|y_d\|^2$. The Assumption (7.57) is not required to obtain an error estimate. A simple and straightforward calculation using the triangle and Cauchy–Schwarz inequality with $(1,\dots,1)^T \in \mathbb{R}^L$ shows

$$\begin{aligned}\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] &\leq \mathbb{E}\left[\left(\sum_{\ell=1}^{L}\|\widehat{\delta}_\ell - \delta_\ell\|\right)^2\right] \leq \left(\sum_{\ell=1}^{L}1^2\right)\mathbb{E}\left[\sum_{\ell=1}^{L}\|\widehat{\delta}_\ell - \delta_\ell\|^2\right]\\ &\leq L\sum_{\ell=1}^{L}\mathbb{E}\Big[\|\widehat{\delta}_\ell - \delta_\ell\|^2\Big].\end{aligned} \tag{7.62}$$

We are then able to bound these differences similarly to the proof of Lemma 7.20 and obtain

$$\mathbb{E}\big[\|\widehat{u}_L^{\mathrm{MLC}} - u_L\|^2\big] \leq cL\left(\sum_{\ell=1}^{L}\frac{2^{-2\gamma\ell}}{m_\ell}\prod_{j=\ell+1}^{L}\big[1 + c2^{-2\gamma j}\big]\right)\|y_d\|^2 \leq cL\left(\sum_{\ell=1}^{L}\frac{2^{-2\gamma\ell}}{m_\ell}\right)\|y_d\|^2.$$

This error is asymptotically larger due to the additional factor L and worsens the final complexity by a logarithmic factor, however, a close inspection of the proof of Theorem 7.21 and especially Lemma 7.18 shows that using this weaker bound simplifies the analysis. This might be of interest if the function J in (7.1) is more complicated or the solution operator S is non–linear.

Similar to the MC estimator we use the CG–method to iteratively solve (7.36) sequentially for $\ell \in \{1,\dots,L\}$. Therefore we define the CG–approximation of $\widehat{\delta}_\ell$ as $\widehat{\delta}_\ell^{\mathrm{CG}}$ and the overall solution on level L as

$$\widehat{u}_L^{\mathrm{MLC,\,CG}} := \sum_{\ell=1}^{L}\widehat{\delta}_\ell^{\mathrm{CG}}.$$

We now derive the main complexity result of this chapter showing improved complexity of MLC similarly to the MLMC complexity Theorem 3.49.

Theorem 7.22 (Complexity of MLC)**.** Let $\alpha > 0$, Assumption 7.2 and Assumption 7.6 be true. Let the cost increase for both an application of S_ℓ and its adjoint S_ℓ^* be at most geometric

$$\mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*] \leq c 2^{\gamma_{\text{Cost}}\ell} \qquad \text{for all } \ell \in \{1, \dots, L\}. \tag{7.63}$$

Then for all $\varepsilon \in (0, 1/e]$ there exists L and $m_1, \dots, m_L$ such that $\mathbb{E}\Big[\|\widehat{u}_L^{\text{MLC, CG}} - u\|^2\Big] \leq \varepsilon^2$ with expected costs bounded by

$$\mathbb{W}\Big[\widehat{u}_L^{\text{MLC, CG}}\Big] \leq c|\log(\varepsilon)| \begin{cases} \varepsilon^{-2}, & \text{if } 2\gamma > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}\log(\varepsilon)^2, & \text{if } 2\gamma = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-2\gamma}{\gamma}}, & \text{if } 2\gamma < \gamma_{\text{Cost}}. \end{cases} \tag{7.64}$$

Proof. The proof is very similar to the proof of Theorem 7.14. We decompose the error

$$\mathbb{E}\Big[\|\widehat{u}_L^{\text{MLC, CG}} - u\|^2\Big] \leq 3\mathbb{E}\Big[\|\widehat{u}_L^{\text{MLC, CG}} - \widehat{u}_L^{\text{MLC}}\|^2\Big] + 3\mathbb{E}\big[\|\widehat{u}_L^{\text{MLC}} - u_L\|^2\big] + 3\|u_L - u\|^2.$$

The bias like term is smaller than $\varepsilon^2/9$ if we choose L as in the proof of Theorem 7.14

$$L \geq \frac{-\log_2(\varepsilon)}{\gamma} + c. \tag{7.65}$$

We now choose the number of samples m_ℓ on level $\ell \in \{1, \dots, L\}$ as follows

$$m_\ell := \phi(\varepsilon) 2^{-\frac{(2\gamma+\gamma_{\text{Cost}})}{2}\ell}, \quad \phi(\varepsilon) := \begin{cases} \varepsilon^{-2}, & \text{if } 2\gamma > \gamma_{\text{Cost}}, \\ \varepsilon^{-2}|\log(\varepsilon)|, & \text{if } 2\gamma = \gamma_{\text{Cost}}, \\ \varepsilon^{-2-\frac{\gamma_{\text{Cost}}-2\gamma}{2\gamma}}, & \text{if } 2\gamma < \gamma_{\text{Cost}}. \end{cases} \tag{7.66}$$

We are able to apply Theorem 7.21, since the additional sample assumption (7.57) is satisfied. To show this, we use the properties of the geometric sum and (7.65)

$$\sum_{\ell=1}^{L} \frac{1}{m_\ell} = \frac{1}{\phi(\varepsilon)} \sum_{\ell=1}^{L} 2^{\frac{2\gamma+\gamma_{\text{Cost}}}{2}\ell} \leq c \frac{2^{\frac{(2\gamma+\gamma_{\text{Cost}})}{2}(L+1)}}{\phi(\varepsilon)} \leq c \left\{ \begin{cases} \varepsilon^{1-\frac{\gamma_{\text{Cost}}}{2\gamma}}, & \text{if } 2\gamma > \gamma_{\text{Cost}}, \\ |\log(\varepsilon)|^{-1}, & \text{if } 2\gamma = \gamma_{\text{Cost}}, \\ 1, & \text{if } 2\gamma < \gamma_{\text{Cost}}. \end{cases} \right\} \leq c.$$

We then use (7.59) such that the variance like term satisfies

$$\mathbb{E}\big[\|\widehat{u}_L^{\text{MLC}} - u_L\|^2\big] \leq c \sum_{\ell=1}^{L} \frac{2^{-2\gamma\ell}}{m_\ell} = c \frac{1}{\phi(\varepsilon)} \sum_{\ell=1}^{L} 2^{-2\gamma\ell + \frac{2\gamma+\gamma_{\text{Cost}}}{2}\ell} = c \frac{1}{\phi(\varepsilon)} \sum_{\ell=1}^{L} 2^{\frac{\gamma_{\text{Cost}}-2\gamma}{2}\ell} \leq c\varepsilon^2. \tag{7.67}$$

The last inequality is a calculation very similar to the proof of Theorem 3.31 for all three cases. Finally, we take care of the iteration error, where we use the Cauchy–Schwarz inequality with $(1, \dots, 1)^T \in \mathbb{R}^L$ to conclude

$$\begin{aligned} \mathbb{E}\Big[\|\widehat{u}_L^{\text{MLC, CG}} - \widehat{u}_L^{\text{MLC}}\|^2\Big] &= \mathbb{E}\left[\|\sum_{\ell=1}^{L} (\widehat{\delta}_\ell^{\text{CG}} - \widehat{\delta}_\ell)\|^2\right] \leq \mathbb{E}\left[\left(\sum_{\ell=1}^{L} 1^2\right)\left(\sum_{\ell=1}^{L} \|\widehat{\delta}_\ell^{\text{CG}} - \widehat{\delta}_\ell\|^2\right)\right] \\ &= L \sum_{\ell=1}^{L} \mathbb{E}\Big[\|\widehat{\delta}_\ell^{\text{CG}} - \widehat{\delta}_\ell\|^2\Big]. \end{aligned} \tag{7.68}$$

Notice that Lemma 7.13 also holds for $\widehat{\delta}_\ell$ instead of $\widehat{u}_\ell^{\mathrm{MC}}$, since (7.15) and (7.36) have the same left–hand side operator. We thus conclude that to reach $\|\widehat{\delta}_\ell^{\mathrm{CG}} - \widehat{\delta}_\ell\| \leq c\varepsilon/L$ and thus $\mathbb{E}\Big[\|\widehat{u}_L^{\mathrm{MLC,\,CG}} - \widehat{u}_L^{\mathrm{MLC}}\|^2\Big] \leq \varepsilon^2/9$ the expected number of iterations n_ℓ on level ℓ is bounded from (7.22)

$$\mathbb{E}[n_\ell] \leq c(|\log_2(\varepsilon/L)| + \mathbb{E}\Big[\|\widehat{\delta}_\ell\|^2\Big]^{1/2}). \tag{7.69}$$

We use (7.53), (7.59) with $L-1$, a calculation similarly to (7.67) and $\varepsilon \in (0, 1/e]$ to bound the increment

$$\begin{aligned}\mathbb{E}\Big[\|\widehat{\delta}_\ell\|^2\Big] &\leq c2^{-2\gamma\ell}\big(\|y_d\|^2 + \mathbb{E}\big[\|\widehat{u}_{L-1}^{\mathrm{MLC}} - u_{L-1}\|^2\big]\big) \leq c2^{-2\gamma\ell}\left(1 + \sum_{\ell=1}^{L-1}\frac{2^{-2\gamma\ell}}{m_\ell}\right)\\ &\leq c2^{-2\gamma\ell}\left(1 + \sum_{\ell=1}^{L}\frac{2^{-2\gamma\ell}}{m_\ell}\right) \leq c2^{-2\gamma\ell}\big(1+\varepsilon^2\big) \leq c.\end{aligned}$$

Now use (7.69), that L grows logarithmically w.r.t. ε in (7.65) and again $\varepsilon \in (0, 1/e]$

$$\mathbb{E}[n_\ell] \leq c(|\log_2(\varepsilon)| + |\log_2(L)| + c) \leq c(|\log_2(\varepsilon)| + |\log_2(|\log_2(\varepsilon)|)| + c) \leq c|\log_2(\varepsilon)|.$$

This bound is independent of ℓ and thus the total costs are upper bounded using the geometric cost increase (7.63) and (7.66)

$$\begin{aligned}\mathbb{W}\Big[\widehat{u}_L^{\mathrm{MLC,\,CG}}\Big] &= \mathbb{E}\left[\sum_{\ell=1}^{L} n_\ell m_\ell(\mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*])\right] \leq c\sum_{\ell=1}^{L}\mathbb{E}[n_\ell]m_\ell 2^{\gamma_{\mathrm{Cost}}\ell}\\ &\leq c|\log_2(\varepsilon)|\phi(\varepsilon)\sum_{\ell=1}^{L} 2^{\frac{(\gamma_{\mathrm{Cost}}-2\gamma)}{2}\ell}.\end{aligned}$$

A straightforward calculation similar to the proof of Theorem 3.31 then shows the complexity result (7.64). As a final remark we note that rounding never increases the asymptotic complexity, since the additional cost of rounding is of order at most $|\log_2(\varepsilon)|\varepsilon^{-\gamma_{\mathrm{Cost}}/\gamma}$. This is not surprising, since the rounding costs never dominate for $2\gamma_{\mathrm{Bias}} = \gamma_{\mathrm{Var}}$ in (3.46) and we have $2\gamma = 2\gamma_{\mathrm{Bias}} = \gamma_{\mathrm{Var}}$. □

We conclude this section by remarking that MLC achieves the optimal complexity (up to logarithmic factors) of $|\log_2(\varepsilon)|\varepsilon^{-2}$ if the variance reduction is larger than the cost increase $2\gamma > \gamma_{\mathrm{Cost}}$. This is in contrast to the standard MC estimator, which only achieves this complexity for the exact sampling case $\gamma_{\mathrm{Cost}} = 0$.

Remark 7.23 (Removing the logarithmic factor)**.** Similar to Remark 7.15 with the same basic idea applied to $\widehat{\delta}_\ell^{\mathrm{CG}}$ for all $\ell \in \{1, \dots, L\}$, we believe that we are able to remove the logarithmic factor $|\log_2(\varepsilon)|$. Therefore, instead of requiring the expected number of iterations to be equal to (7.69), we w.l.o.g. may assume that

$$\mathbb{E}[n_\ell] \leq c(|\log_2(1/L)| + \mathbb{E}\Big[\|\widehat{\delta}_\ell\|^2\Big]^{1/2}) \leq c(1 + |\log_2(L)|) \leq c|\log_2(|\log_2(\varepsilon)|)|,$$

where we used that L is logarithmic w.r.t. ε in (7.65). We thus have replaced the logarithmic factor $|\log_2(\varepsilon)|$ with the much smaller $|\log_2(|\log_2(\varepsilon)|)|$. This additional factor is a consequence of (7.68), where we applied the Cauchy–Schwarz inequality. We believe that this leads to an estimate that is not sharp and thus to an additional factor L similarly

to (7.62). The latter is itself is not sharp compared to the improved but technically more challenging (7.59) without the extra L. If a similar improvement can be derived for (7.68), then the additional factor $|\log_2(|\log_2(\varepsilon)|)|$ can entirely be removed. For most practical purposes the factor $|\log_2(|\log_2(\varepsilon)|)|$ grows so slowly that it may be viewed as a constant not worth the additional effort to remove it. ◇

7.4 Numerical experiments

We numerically verify some of the results stated in this chapter. As constraint we use the elliptic boundary value problem from Section 2.3

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

where the diffusion coefficient $a := \exp(\kappa)$ is lognormal. Here $\kappa \sim N(0,\mathcal{C})$, where $\mathcal{C}$ is the covariance operator with the Whittle–Matérn kernel with smoothness $\nu := 3/2$, variance $\sigma^2 := 1$ and correlation length $\ell := 0.5$. We furthermore use the Tikhonov regularization parameter $\alpha := 1$. We uniformly refine the finite element mesh starting at 9 nodes up to level $L := 8$ with 16641 nodes. We compute reference solutions u^{Ref} on level $\ell = 1,\dots,7$ with a sufficiently large number of samples using the standard MC method. We define and approximate the bias as follows

$$\operatorname{Bias}(u_\ell) := \|u_\ell - u\| \approx \|u_\ell^{\mathrm{Ref}} - u_7^{\mathrm{Ref}}\| \quad \text{for all } \ell \in \{1,\dots,6\}$$

and extrapolate it for level $\ell = 7, 8$ such that $\operatorname{Bias}(u_\ell) := \operatorname{Bias}(u_{\ell-1})/4$. The sample allocation for both the MC and MLC estimator is not straightforward to determine. However, a computation using (7.10) and $\alpha = 1$ shows

$$u = \mathbb{E}[S^*(y_d - Su)]. \tag{7.70}$$

Here the control appears on the left and right–hand side. We eliminate the appearance on the right–hand side to obtain the approximation

$$u_\ell \approx \mathbb{E}[S_\ell^*(y_d - S_\ell u)]. \tag{7.71}$$

The MC estimator of this quantity with a single sample assuming a fixed u then gives the covariance for $\ell, j \in \{1,\dots,L\}$

$$C_{\ell,j} := \mathbb{E}\big[(S_\ell^*(y_d - S_\ell u) - \mathbb{E}[S_\ell^*(y_d - S_\ell u)], S_j^*(y_d - S_j u) - \mathbb{E}\big[S_j^*(y_d - S_j u)\big])_{L^2}\big]. \tag{7.72}$$

We use this as ad hoc definition for the covariance matrix which was computed using $N := 10^4$ pilot samples. Formally, for any orthonormal basis $(\psi_n)_{n\in\mathbb{N}}$ of $L^2(D)$ we collapse the covariance between two random fields into a single value

$$C_{\ell,j} = \sum_{n=1}^{\infty} \mathbb{C}\mathrm{ov}\big[S_\ell^*(y_d - S_\ell u), S_j^*(y_d - S_j u)\big](\psi_n, \psi_n).$$

We also use the pilot samples to verify the cost assumption

$$w_\ell := \mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*] \le c 2^{\gamma_{\mathrm{Cost}}\ell} \quad \text{for all } \ell \in \{1,\dots,L\}. \tag{7.73}$$

We plot the bias, the variance $C_{\ell,\ell}$, the variance difference

$$\mathbb{V}\left[\widehat{u}_\ell^{\mathrm{MC}} - \widehat{u}_{\ell-1}^{\mathrm{MC}}\right] \approx C_{\ell,\ell} + C_{\ell-1,\ell-1} - 2C_{\ell-1,\ell} \tag{7.74}$$

as well as the costs in Figure 7.1. We see that the assumption of $\gamma_{\mathrm{Cost}} = 2$ is satisfied, whereas the bias rate is close to $\gamma = 2$ and the variance difference rate is slightly smaller than $2\gamma = 4$. Hence it is safe to assume that $\gamma_{\mathrm{Cost}} < 2\gamma$ and that the best complexity case of ε^{-2} (ignoring logarithmic factors) holds for the MLC estimator. We conclude that the

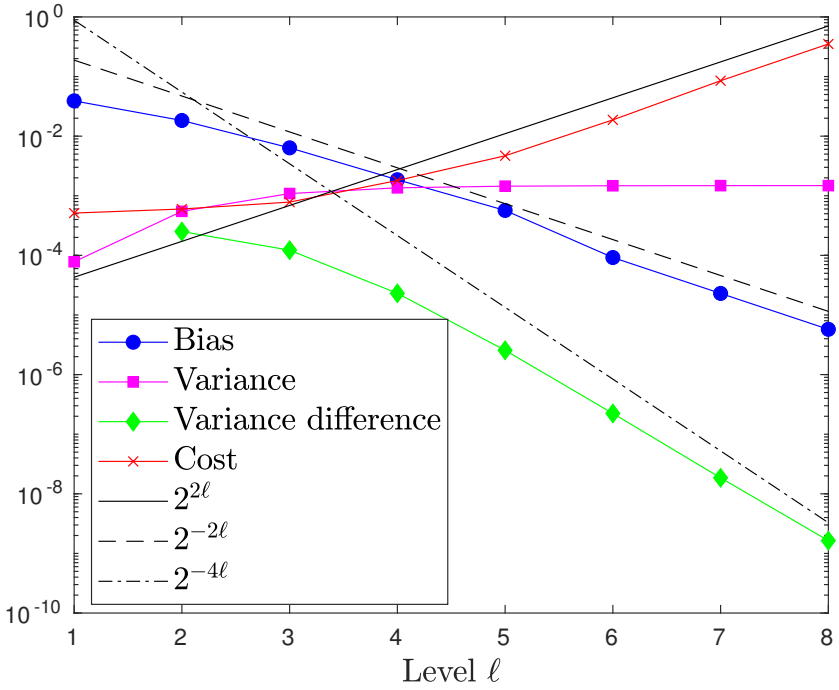

Figure 7.1: Ad hoc defined bias, variance $C_{\ell,\ell}$, variance difference (7.74) of the right–hand side in (7.71) and costs $\mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*]$ for all levels $\ell \in \{1, \dots, L\}$. The reference rates are drawn black.

cost assumption (7.73) as well as Assumption 7.6 is satisfied. This is not surprising, since both Assumption 7.2 and Assumption 7.6 are satisfied for the elliptic PDE constraint due to Theorem 2.34 respectively Theorem 2.39 and Theorem 2.40.
We define the *idealized MSE* to match the form of the standard MLMC estimator as follows

$$\mathrm{IMSE}(\widehat{u}_\ell) := \mathrm{Bias}(u_\ell)^2 + \sum_{\ell=1}^{L} \frac{C_{\ell,\ell} + C_{\ell-1,\ell-1} - 2C_{\ell,\ell-1}}{m_\ell}.$$

Here the nominator of the fractions approximate $\mathbb{V}\left[\widehat{u}_\ell^{\mathrm{MC}} - \widehat{u}_{\ell-1}^{\mathrm{MC}}\right]$ due to (7.74). We now require the variance part to be equal to the square of the bias on level ℓ, that is

$$\sum_{\ell=1}^{L} \frac{C_{\ell,\ell} + C_{\ell-1,\ell-1} - 2C_{\ell,\ell-1}}{m_\ell} = \mathrm{Bias}(u_L)^2.$$

The standard MLMC sample allocation, see Theorem 3.46, is then used to compute $m_1, \dots, m_L$ to minimize the cost. The sample allocation for the MC estimator is computed similarly with $C_{L,L}/m_L$ as variance part. The allocations are listed in Table 7.1. Here

we remark that the sample allocation algorithm showed that MLC only reduces the costs if the level is greater or equal to 4, otherwise the standard MC estimator is cheaper. Note that we do not account for the iteration costs of solving linear systems. Instead, we apply the CG–method until the residual has the norm 10^{-9}, which exceeds the range of the MSE that we care about. This removes the logarithmic factor in the MC complexity Theorem 7.14 and MLC complexity Theorem 7.22. In practice, we observed that we require three iterations of the CG–method to reach this accuracy, which is fast and a result of the rather large $\alpha = 1$. We thus have approximately four state and adjoint solves, one for the right–hand side and three for the solution of the linear system. This allows us to simply multiply the cost by a factor of four. The computed costs are then defined as

$$4 \sum_{\ell=1}^{L} m_\ell (\mathbb{W}[S_\ell] + \mathbb{W}[S_\ell^*]).$$

We plot the computed costs versus the idealized MSE for both MC and MLC in Figure 7.2. The idealized MSE underestimates the true MSE, since we did not account for the bias that is obtained by sampling

$$\mathbb{E}\big[\widehat{u}_\ell^{\mathrm{MC}}\big] \neq u_\ell.$$

and thus have

$$\mathbb{E}\big[\|\widehat{u}_\ell^{\mathrm{MC}} - u_\ell\|^2\big] \neq \mathbb{V}\big[\widehat{u}_\ell^{\mathrm{MC}}\big].$$

Remark 7.12 shows that the bias decays faster than the variance. A similar result should also hold for the MLC estimator and thus the idealized MSE should be a reasonable metric that can be used for the sample allocation.
We account for this by plotting the computed costs versus the actual MSE in the left image of Figure 7.3. We compute the MSE with the help of the reference solution on level $L = 7$ using 100 independent runs for both the MC and MLC estimator

$$\mathrm{MSE}(\widehat{u}_\ell) \approx \frac{1}{100} \sum_{i=1}^{100} \|\widehat{u}_\ell^i - u_7^{\mathrm{Ref}}\|^2. \tag{7.75}$$

The reference solution on level 7 allows us to obtain numerical results only up to level 6. This is rather unsatisfactory, however, computing a reference solution on level $L = 8$ with MC is prohibitively expensive. In fact, the costs to compute the MC estimator on level 8 is approximately two years, whereas the MLC estimator can be computed in less than three days. Therefore, we run another experiment using a level 8 reference solution obtained from the MLC estimator, which we call u_8^{Ref}. This additional reference solution allows us to compute the MSE on level 7 for the MLC estimator. The right plot of Figure 7.3 shows these values. We conclude that the MLC estimator is a significant improvement over the standard MC estimator. In particular, the cost of the MC estimator follows the predicted costs of ε^{-3} in Theorem 7.14 to reach a MSE of ε^{-2}, whereas the MLC estimator seems to have the optimal cost rate of ε^{-2} in Theorem 7.22. Once again, we iterate the CG–algorithm a fixed number of times and thus we do not measure the logarithmic factors. We believe that this rate is visible rather late, since the constant in (7.56) is

$$\prod_{j=\ell+1}^{L} \left[1 + c\left(\frac{1}{m_j} + 2^{-\gamma j} + 2^{-2\gamma j}\right)\right].$$

The increase of this constant compared to a previous level is rather large for small levels L since $m_1, \dots, m_L$ are typically small and $2^{-\gamma L}$ is comparatively large.

Level	$\widehat{u}_2^{\mathrm{MC}}$	$\widehat{u}_3^{\mathrm{MC}}$	$\widehat{u}_4^{\mathrm{MC}}$	$\widehat{u}_4^{\mathrm{MLC}}$	$\widehat{u}_5^{\mathrm{MC}}$	$\widehat{u}_5^{\mathrm{MLC}}$
2	2	-	-	-	-	-
3	-	27	-	401	-	4,749
4	-	-	395	33	-	381
5	-	-	-	-	4,495	80
Cost	0.0048s	0.08s	2.8s	1.6s	84s	21s
Level	$\widehat{u}_6^{\mathrm{MC}}$	$\widehat{u}_6^{\mathrm{MLC}}$	$\widehat{u}_7^{\mathrm{MC}}$	$\widehat{u}_7^{\mathrm{MLC}}$	$\widehat{u}_8^{\mathrm{MC}}$	$\widehat{u}_8^{\mathrm{MLC}}$
3	-	190,864	-	3,151,628	-	51,381,086
4	-	15,289	-	252,453	-	4,115,744
5	-	3,211	-	53,018	-	864,354
6	173,982	499	-	8,238	-	134,290
7	-	-	2,795,608	1,130	-	18,411
8	-	-	-	-	44,777,906	2,667
Cost	3.6h	0.25h	264h	4.2h	729d	2.9d

Table 7.1: Sample allocation for the MC and MLC estimator on different levels. The entries correspond to the number of samples the estimator uses on the respective level. The costs correspond to the computed costs of the estimator, that is all samples of all levels assuming three iterations of the CG–algorithm plus the computation of the right–hand side. The costs are given in seconds (s), hours (h) or days (d). The estimator $\widehat{u}_1^{\mathrm{MC}}$ only requires a single sample and costs approximately 0.0021 seconds. The estimators $\widehat{u}_2^{\mathrm{MLC}}$ and $\widehat{u}_3^{\mathrm{MLC}}$ have larger variance than the corresponding MC estimators and are thus not listed.

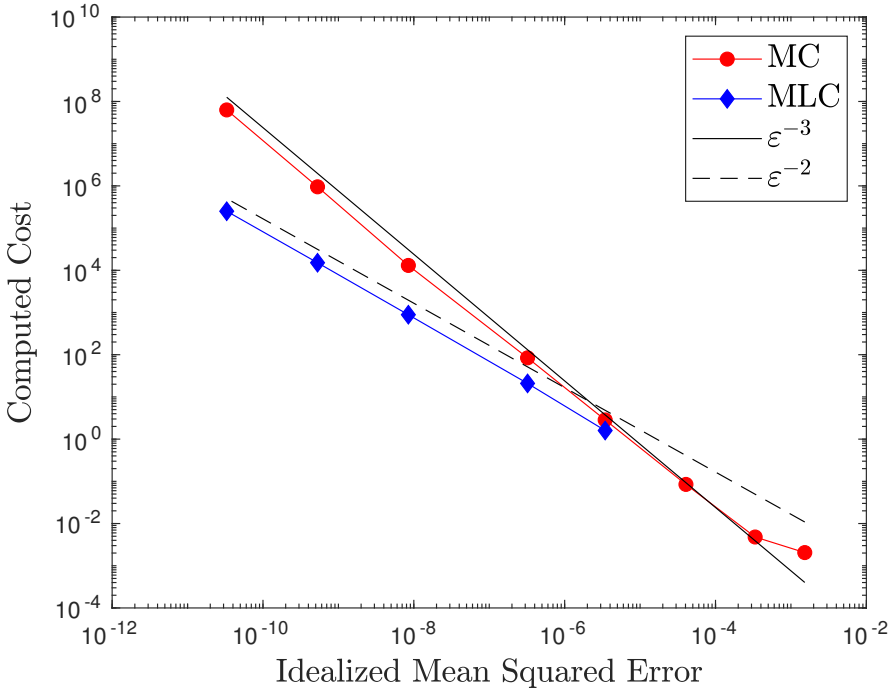

Figure 7.2: Computed costs versus the idealized MSE of both the MC and MLC estimator on all levels. The idealized MSE does not account for any bias obtained from sampling and thus has to understood as a lower bound for the true MSE.

	1	2	3	4	5	6	7
MC	0.54	0.93	1.15	1.25	1.19	2.24	-
MLC	-	-	-	1.07	1.17	2.16	2.18

Table 7.2: Quotient of the true MSE divided by the idealized MSE for the MC and MLC estimator on different levels.

Finally, we plot the true costs of our implementation w.r.t. the MSE in Figure 7.4. We used a computer with four equally fast processors and thus the cost is cut by roughly a quarter, i.e. 0.98 hours actual time used for a sample of $\widehat{u}_7^{\mathrm{MLC}}$ versus the computed costs of approximately 4.2 hours. This observation seems to be valid for datapoints on fine levels but not for coarse levels. We believe that this is a result of our implementation which is not very efficient for a small number of samples. In particular, we think that we have sample independent fix costs of at least two seconds due to the parallelization every time we solve a linear system. This affects the performance of the coarse estimators $\widehat{u}_1^{\mathrm{MC}}, \dots, \widehat{u}_4^{\mathrm{MC}}$ such that they have almost the same costs. The $\widehat{u}_4^{\mathrm{MC}}$ is in fact cheaper than $\widehat{u}_4^{\mathrm{MLC}}$ contrary to the computed costs. The latter estimator uses two levels and thus has to solve two linear systems, which means we pay the fix costs twice. The relative overhead gets smaller as we increase the number of samples and thus the estimated costs closely match the true costs on fine levels.
We compare how well the idealized MSE, which we use to compute the sample allocation, approximates the true MSE. The quotient of the true MSE divided by the idealized MSE is given in Table 7.2 and is roughly in the correct range. We believe that the larger values > 2 are a consequence of an inaccurate value of the bias on level 6, which can be seen by the kink from level 5 to 6 in Figure 7.1. Here the bias is smaller by a factor of 6.1 compared to the previous level instead of the expected value of slightly smaller than 4. The bias on level 7 is then extrapolated by dividing the bias on level 6 by 4. Hence, it seems reasonable that the idealized MSE assumes a bias that is too small and thus underestimates the true MSE on level 6 and 7.

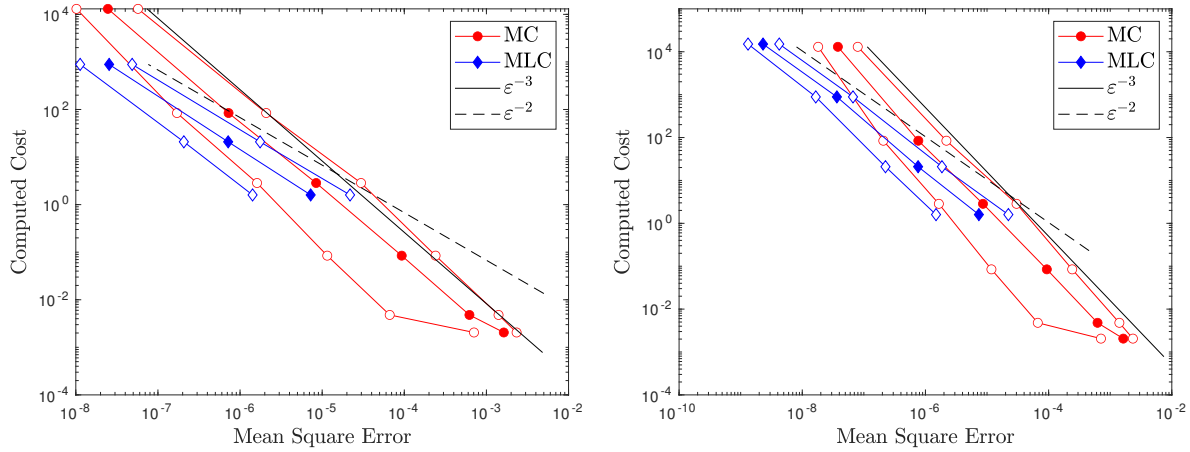

Figure 7.3: The computed cost over 100 runs is plotted with thick markers for both the MC and MLC estimator w.r.t. the MSE (7.75). The left and right lines with white marker are the bounds of the region such that the error of 90 percent of the samples are within this region. The left plot uses the MC reference solution on level 7 and the right plot uses the MLC reference solution on level 8. Both images use the computed costs listed in Table 7.1.

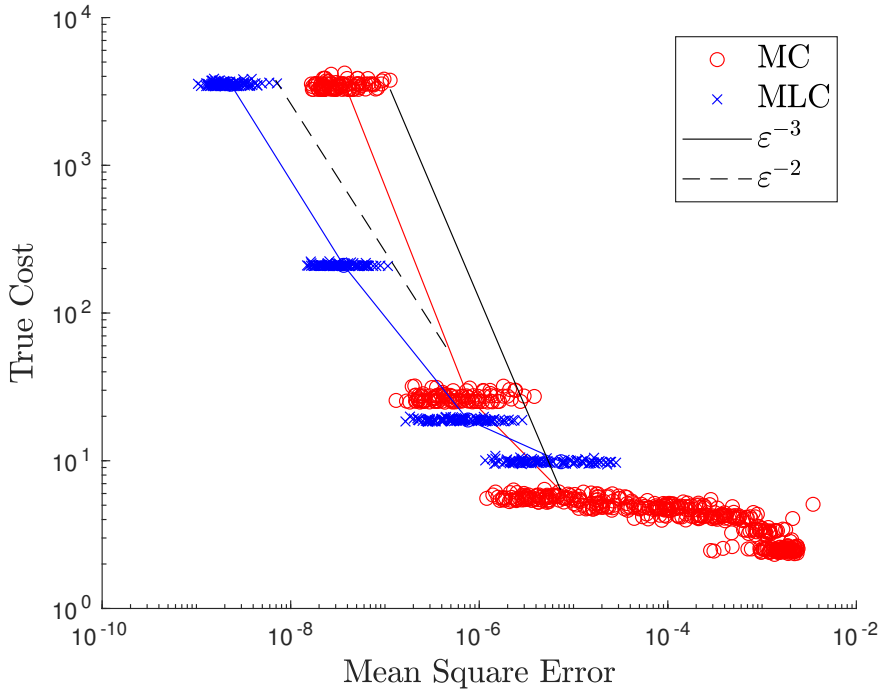

Figure 7.4: True cost versus the MSE of both the MC estimator on level $1, \ldots, 6$ and MLC estimator on level $4, \ldots, 7$. Each one of the 100 datapoints represents the error and cost of one independently computed estimator. The lines show the average costs versus the MSE, reference rates are drawn black. The MSE was computed using the MLC estimator as reference and the true costs are in seconds for an actual implementation using four processors.

Chapter 8
Conclusion and outlook

Conclusion. We briefly summarize the main results of the thesis. We reformulated the estimation of the mean of an output quantity of interest as linear regression problem and used well–known methods to derive the normal equations and the variance minimal BLUE. Afterwards, we introduced a cost constraint and optimized the sample allocation to obtain the SAOB. This estimator has the smallest variance in the class of linear unbiased estimators with a fixed budget. We furthermore derived slightly stronger results by first optimizing the sample allocation and then choosing the coefficients of the respective BLUE. The resulting estimator exists but is not necessarily unique. We showed that there exists a sparse version of it in the sense that at most L model groups are used. We proved that the SAOB achieves optimal asymptotically complexity in the class of linear unbiased estimators. However, an exact expression for its complexity is very difficult to derive, since the number of samples or the coefficients of the linear combination are given implicitly as a solution of an optimization problem. Nevertheless, we introduced the Richardson extrapolation estimators to obtain upper bounds on the costs of the SAOB. We further verified that the asymptotic cost of the SAOB is optimal for linear unbiased estimation.

Finally, we showed that variance reduction techniques can indirectly be applied to the risk neutral optimal control problem. We pushed the MLMC discretization onto the deterministic control which leads to a sequence of convex optimal control problems. We verified that the resulting formulation is well posed and reduces the variance compared to MC. A challenge of this approach was the analysis since the resulting estimators are not unbiased.

The numerical experiments in this thesis verified our results. We showed that the Richardson extrapolation estimators, the SAOB and the MLC estimator are effective variance reduction methods in the context of uncertainty quantification with PDEs.

We now give a short list of open questions, encountered problems and future research directions in the context of this thesis.

Covariance matrix. It is unrealistic to assume that the model covariance matrix C in (2.3) is known but the mean is not. We believe that this is the biggest drawback of the BLUE and thus also of the SAOB. There are several possible approaches to obtain an approximation of the covariance matrix, however, it is not clear how large the error of such an estimator is.

The simplest solution is to use pilot samples to compute the sample covariance. These pilot samples are then not used for the estimator. This approach clearly increases the cost, however, we believe that in some scenarios this does not effect the asymptotic complexity. Assume that the complexity of a single evaluation of all models satisfies

$$\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}} \leq \varepsilon^{-2+\delta}$$

for some $\delta > 0$, i.e. the evaluation of a sample of $(Z_1, \dots, Z_L)$ is slightly cheaper than the optimal complexity ε^{-2}. Then we are able to spend an increasing amount of the budget

of order $\varepsilon^{-\delta}$ without worsening the asymptotic cost, that is

$$\text{Cost} = \underbrace{\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}}_{\text{Cost for rounding}} + \underbrace{\varepsilon^{-2}}_{\text{Cost for the estimator}} + \underbrace{N_{\text{pilot}}\varepsilon^{-\gamma_{\text{Cost}}/\gamma_{\text{Bias}}}}_{\text{Cost for pilot samples}} \leq c\varepsilon^{-2},$$

where $N_{\text{pilot}} = \varepsilon^{-\delta}$ is the number of pilot samples. This means that asymptotically the number of pilot samples may be increased to infinity without worsening the asymptotic complexity. Furthermore, if we choose $N_{\text{pilot}} = \varepsilon^{-\delta/2}$, then the cost for the pilot samples compared to the cost for the estimator is negligible. Therefore, if we are able to approximate the covariance matrix with N_{pilot} samples sufficiently well, then the computed SAOB with this matrix converges to the true SAOB. However, it is not straightforward to verify that $N_{\text{pilot}} = \varepsilon^{-\delta}$ is sufficient to show that the sample covariance matrix converges to the true covariance for $\varepsilon \to 0$, since the covariance matrix changes if ε decreases due to additional fine grid models. An error analysis for the SAOB with approximated covariance matrix also has to be conducted to make the above statements mathematically sound. In particular, an error analysis of the problem in (5.22) w.r.t. perturbations in the covariance matrix C^k and costs W^k should be carried out.
The authors of [9] and [112] construct efficient multilevel estimators to compute the variance of a QoI. These approaches may be adapted to compute an even cheaper approximation of the covariance matrix compared to the sample covariance. However, similar to the naive MLMC estimator in Chapter 7, a straightforward MLMC approximation of the entries of the covariance matrix does not ensure that this approximation is positive semi–definite. Hence, the corresponding optimization problem for the SAOB may be ill–posed. Maybe similar methods to the one outlined in Chapter 7 can be used to push the MLMC estimator to the coefficients β of the BLUE instead of the covariance matrix.
Aside from the asymptotic cost increase of the estimator, it is not clear how big the constants in front of the costs for the pilot samples are. In particular, if few samples are used, then the estimator may be overconfident, and we might assume that the correlation is much higher than it truly is. This then leads to a bad estimator. It is of course possible to avoid pilot samples or to reuse the pilot samples for the estimation, however, this introduces a bias. Whether this bias is actually significant and has to be avoided or can be fully ignored is not clear as of the writing of the thesis. This may also depend on the particular estimation problem. We also did not conduct any conclusive numerical experiments. If an unbiased estimator is desired, then clearly we are able to use the estimator

$$\widehat{\mu}_\alpha := \widehat{\mu}^1_{\alpha/2}[\widehat{C}^2] + \widehat{\mu}^2_{\alpha/2}[\widehat{C}^1].$$

Here the estimator $\widehat{\mu}^1_{\alpha/2}[\widehat{C}^2]$ is the SAOB that uses the sample covariance matrix $\widehat{C}^2$ obtained from the samples of $\widehat{\mu}^2_{\alpha/2}$ to compute its coefficients β and vice versa. This approach ensures that the resulting estimator is unbiased, however it is not clear if and by how much the variance increases.

Condition number of the covariance matrix. The condition of the covariance matrix C is typically large. If the quantities Z_L converge to the exact model Z, then the covariance matrix C has, at least asymptotically, almost constant entries with values approximately equal to $\mathbb{V}[Z]$. Hence, there is an eigenvalue close to $L\mathbb{V}[Z]$ with eigenvector consisting of entries that are all close to one. On the other hand, a variance reduction can often be achieved by obtaining a small value of

$$\beta^T C \beta = \mathbb{V}\left[\sum_{\ell=1}^{L} \beta_\ell Z_\ell\right] \approx \lambda_{\min}$$

with $\|\beta\| = 1$ and $\lambda_{\min}$ denoting the smallest eigenvalue. Therefore, the larger the achievable variance reduction, the smaller the smallest eigenvalue. Since the value of $\mathbb{V}[Z_\ell - Z_{\ell-1}]$ converges to zero, $\lambda_{\min}$ does so as well. Hence, the condition of C diverges to $+\infty$. Consequently, the better the variance reduction, the worse the condition of C. We thus believe that the algorithms that depend on C^{-1} or some $(C^\ell)^{-1}$ may be unstable or yield inaccurate results. This undesirable property also occurs for some Richardson extrapolation methods. We verified that the coefficients of the SAOB converge to the coefficients of the RE estimator in specific circumstances and thus we expect the SAOB may suffer from the same inaccuracies. The significance of this problem can be reduced if we fix the coupling number κ to a small value or we carefully choose a small subset of all $2^L - 1$ model groups. However, this results in a potentially decreased variance reduction which may or may not be significant.

Hybrid estimators. The SAOB does not use any structural assumptions about the QoI. In particular, the SAOB only requires the covariance matrix and model costs to achieve a variance reduction. However, this generality comes at the cost of requiring an estimate of the covariance matrix. The numerical experiments in Chapter 6 have shown that the improvement of the SAOB over the MLMC and RE estimators is small if we are in the asymptotic regime. In contrast, far away from this regime the SAOB, or in general any BLUE, is able to provide significant improvements. Therefore, it may make sense to construct a hybrid estimator as follows

$$\widehat{\mu}_L^{\mathrm{MLMC}+\mathrm{SAOB}} := \sum_{\ell=L_{\mathrm{coarse}}+1}^{L} \frac{1}{m_\ell} \sum_{i=1}^{m_\ell} \left(Z_\ell^{i,\ell} - Z_{\ell-1}^{i,\ell} \right) + \widehat{\mu}_{L_{\mathrm{coarse}}}^{\mathrm{SAOB}}.$$

The SAOB here only estimates $\mu_{L_{\mathrm{coarse}}}$ and thus only needs the covariance of the coarse models $Z_1, \dots, Z_{L_{\mathrm{coarse}}}$, which is in general cheap to obtain. At the same time, we use the MLMC estimator for the high fidelity models, which does not require the full covariance matrix. We thus use the SAOB in the regime of coarse models where it has the biggest advantage over other methods. An open question is of course, that it is not clear how to choose L_{coarse} or how to increase L_{coarse} if we want to increase the accuracy of the estimator.

Regularization. Another approach would be to include some prior information for the coefficients. As an example, we might add a regularization term to (5.22) in form of the MLMC estimator

$$\begin{aligned} &\min_{\beta} \sum_{k=1}^{K} \left[(\beta^k)^T C^k \beta^k W^k \right]^{1/2} + \alpha_{\mathrm{reg}} \sum_{k=1}^{L} \left[(\beta^k - \beta^{\mathrm{MLMC},k})^T C^k (\beta^k - \beta^{\mathrm{MLMC},k}) W^k \right)\Big]^{1/2} \\ &\text{such that} \qquad \sum_{k=1}^{K} P^k \beta^k = \alpha. \end{aligned}$$

Here α_{reg} is a regularization parameter, $\beta^{\mathrm{MLMC},k}$ the coefficients and $S^1, \dots, S^L$ the model groups of the MLMC estimator. Depending on the size of α_{reg} we may be close to a MLMC estimator or far away from it. In particular, the parameter α_{reg} is allowed to contain some information about the accuracy of the estimate for the covariance matrix. Ideally, if we do not know much about the covariance matrix, which may be derived from error estimates, the value α_{reg} should be large and conversely, if the estimate of the covariance is very accurate, the value of α_{reg} should be small or zero. This approach ensures that the

potentially inaccurate covariance matrix computed from only a few samples does not lead to an inaccurate estimator.

Efficient algorithm to solve the sample allocation problem. The sample allocation problem in Section 5.2 to compute m, or the equivalent problem to compute β in Section 5.3 are both convex optimization problems. However, the number of variables depends exponentially on the number of levels L if we do not restrict the coupling of the models, that is $\kappa = +\infty$. Therefore, it is not clear whether there is a polynomial time algorithm (polynomial in L) that solves this problem. Nevertheless, the optimization problem (5.22) exhibits some interesting local behaviour. Let us assume that some iterative algorithm returns a suboptimal allocation that uses the model groups $S^1, \dots, S^L$ and we want to check whether it makes sense to use the model group S^{L+1}. Then if for all β^{L+1} the minimization problem

$$\begin{aligned} &\min_{\beta^1,\dots,\beta^L} J(\beta) := \sum_{\ell=1}^{L} \left(W^\ell (\beta^\ell)^T C^\ell \beta^\ell\right)^{1/2} \\ \text{such that} \qquad &\sum_{\ell=1}^{L} P^\ell \beta^\ell = P^{L+1}\beta^{L+1} \end{aligned} \tag{8.1}$$

has a minimum smaller than $\left(W^{L+1}(\beta^{L+1})^T C^{L+1}\beta^{L+1}\right)^{1/2}$ the SAOB does not use the model group S^{L+1}. This can be proven by contradiction, thus assume that the SAOB does and write down this estimator as follows

$$\widehat{\mu}_\alpha^{\text{SAOB}} = \widehat{\mu} + \frac{1}{m_{L+1}} \sum_{i=1}^{m_{L+1}} \sum_{\ell \in S^{L+1}} \beta_{[\ell]}^{L+1} Z_\ell^{i,L+1}, \tag{8.2}$$

where $\widehat{\mu}$ is a suitable estimator. Now, the second part of this estimator is an unbiased estimator for $(P^{L+1}\beta^{L+1})^T\mu$, however, the function J in (8.1) is the variance of another SAOB with, by assumption, smaller variance then the estimator using only S^{L+1}. Due to the independence structure, we can replace the right estimator of (8.2) with the SAOB obtained from (8.1) leading to a smaller variance and thus a contradiction. The scale invariance now implies that the statement we made so far is equivalent to show that the max-min problem

$$\begin{aligned} &\max_{\|\beta^{L+1}\|=1} \min_{\beta^1,\dots,\beta^L} J(\beta) := \sum_{\ell=1}^{L} \left(W^\ell (\beta^\ell)^T C^\ell \beta^\ell\right)^{1/2} - \left(W^{L+1}(\beta^{L+1})^T C^{L+1}\beta^{L+1}\right)^{1/2} \\ \text{such that} \qquad &\sum_{\ell=1}^{L} P^\ell \beta^\ell = P^{L+1}\beta^{L+1} \end{aligned} \tag{8.3}$$

has a maximizer strictly smaller than 0. As an example why this result might be useful, assume that the MLMC estimator for μ_L has a smaller variance than the MC estimator for μ_L using the same budget. Then, the SAOB never uses the model group that only contains Z_L. A possible research direction is to use this local property to exclude some model groups that are not useful to reduce the variance. Then the optimization problem may have a smaller number of variables instead of exponentially many.

Flat optimization surface. The numerical experiments in Section 6.5 for the true costs showed that the variance of the compared estimators (except for MC) are all very

similar. This is not surprising, since the variance reduction ensures that most of the costs are on the coarse grids allowing a different number of fine grid evaluations, or use of different model groups without worsening the overall costs too much. More specifically, using the RE coefficients or coefficients close to it, the variance corresponding to the model group S^L with the high fidelity model is

$$(\beta^L)^T C^L \beta^L W^L \approx c 2^{-(\gamma_{\text{Var}} - \gamma_{\text{Cost}})L}. \tag{8.4}$$

This is asymptotically for $L \to +\infty$ very small if $\gamma_{\text{Cost}} < \gamma_{\text{Var}}$ compared to the value of the model group S^1 which only consists of the coarsest model

$$(\beta^1)^T C^1 \beta^1 W^1 \approx c.$$

For hierarchical models, the cost of the discretization increases geometrically, thus adding or removing the models $1, \dots, L/2$ from the model group S^L, which contains the high fidelity model, is not going to significantly change W^L. Furthermore, adding or removing these models does not change the overall variance that much since (8.4) contributes asymptotically nothing and the variance reduction rate γ_{Var} is often achieved with only a few fine grid levels. Now observe that there are a total of $2^{L/2}$ possible combinations if we keep the fine models $Z_{L/2+1}, \dots, Z_L$ in S^L and add or remove the models $Z_1, \dots, Z_{L/2}$. The result is thus that there are exponentially many directions that lead to an almost identical function value. However, this means that the function J we optimize over in (5.22) is flat at the minimum. This often poses difficulties for standard numerical optimization methods since the gradient becomes small and the Hessian ill–conditioned. This can be mitigated by choosing to optimize only over a subset of all $2^L - 1$ model groups and thus removing some flat direction. If a minimizer is found, then suitably adding or removing some model groups could be used to obtain the global minimizer.

Extension to non–linear estimators. The author of [105] used the MFMC estimator with a low fidelity model that is adaptively improved to reduce the variance. A similar approach might also be viable for the BLUE. The MFMC estimator requires models that are cheap and correlate well with the high fidelity model. Let us assume that $\mathbb{V}[Z] := \mathbb{V}[Z_\ell(\theta)] := 1$ where $Z_\ell(\theta)$ is the low fidelity model and θ some model parameter. Then we want to choose θ to minimize

$$\begin{aligned} 1 - \rho^2 = 1 - \mathbb{C}\text{ov}[Z, Z_\ell(\theta)]^2 &= (1 - \mathbb{C}\text{ov}[Z, Z_\ell(\theta)])(1 + \mathbb{C}\text{ov}[Z, Z_\ell(\theta)]) \\ &\leq 2(\mathbb{C}\text{ov}[Z, Z - Z_\ell(\theta)] \leq 2(\mathbb{V}[Z - Z_\ell(\theta)])^{1/2}. \end{aligned}$$

It is now straightforward to generalize this idea for the SAOB with multiple low fidelity models. The optimization problem (5.22) then reads

$$\begin{aligned} \min_{\beta, \theta} \quad & \sum_{k=1}^{K} \left((\beta^k)^T C^k(\theta) \beta^k W^k(\theta) \right)^{1/2} \\ \text{such that} \quad & \sum_{k=1}^{K} P^k \beta^k = e_L. \end{aligned} \tag{8.5}$$

Here the costs as well as the covariance matrix now depend on θ

$$C_{\ell,j}(\theta) = \mathbb{C}\text{ov}[Z_\ell(\theta), Z_j(\theta)].$$

The quantities $Z_1, \dots, Z_L$ are potentially combined in a non–linear way since the cost function is

$$\sum_{k=1}^{K} \left((\beta^k)^T C^k(\theta) \beta^k W^k(\theta) \right)^{1/2} = \sum_{k=1}^{K} \left(\mathbb{V}\left[\sum_{\ell \in S^k} \beta^k_{[\ell]} Z_\ell(\theta) \right] W^k(\theta) \right)^{1/2}.$$

The pilot samples, which we use to compute the sample covariance matrix, can now be used to change the low fidelity models to reduce the variance. Let us assume that (8.5) has a well–defined minimizer (β^*, θ^*). Then the SAOB with parameters (β^*, θ^*) achieves the smallest variance in the class of unbiased estimators for μ_L that have equal or smaller cost, depend linearly on β and non–linearly on θ and only use the QoIs $Z_1(\theta), \dots, Z_L(\theta)$. In this sense, the SAOB is still variance minimal and thus a sensible choice even if non–linearities arise. Furthermore, similar to the standard SAOB, once the models $Z_1(\theta^*), \dots, Z_L(\theta^*)$ are fixed we require at most L model groups. A disadvantage may be that solving (8.5) is difficult and impractical. Nevertheless, such an approach includes non–linear surrogate models like Gaussian processes, support vector machines or neural networks.

Variance reduction for root-finding algorithms. The variance reduction technique for the optimal control problem in Chapter 7 can be generalized to root-finding algorithms. Assume that we want to find a deterministic $\theta \in \Theta$ such that for a parametric random variable $Z : \Theta \times \Omega \to X$

$$\mathbb{E}[Z(\theta)] = 0.$$

The discretized condition with $Z_1, \dots, Z_L$ then reads

$$\mathbb{E}[Z_\ell(\theta_\ell)] = 0 \qquad \text{for all } \ell \in \{1, \dots, L\}.$$

We define $\delta_\ell := \theta_\ell - \theta_{\ell-1}$ and we linearise the left–hand side of the equation

$$\mathbb{E}[Z_\ell(\theta_\ell)] = \mathbb{E}[Z_\ell(\theta_{\ell-1} + \delta_\ell)] = \mathbb{E}[Z_\ell(\theta_{\ell-1})] + \mathbb{E}[\nabla_\theta Z_\ell(\theta_{\ell-1})]^T \delta_\ell + o(\|\delta_\ell\|).$$

A single step of the Newton method, that is ignoring $o(\|\delta_\ell\|)$, and assuming that $\theta_{\ell-1}$ is known, solves for the Newton step s^1_ℓ with the starting iterate $\theta_{\ell-1}$

$$\mathbb{E}[\nabla_\theta Z_\ell(\theta_{\ell-1})]^T s^1_\ell = -\mathbb{E}[Z_\ell(\theta_{\ell-1})] = 0 - \mathbb{E}[Z_\ell(\theta_{\ell-1})] = \mathbb{E}[Z_{\ell-1}(\theta_{\ell-1}) - Z_\ell(\theta_{\ell-1})]. \quad (8.6)$$

The next iteration of the Newton method then has to solve

$$\mathbb{E}\big[\nabla_\theta Z_\ell(\theta_{\ell-1} + s^1_\ell)\big]^T s^2_\ell = \mathbb{E}\big[Z_{\ell-1}(\theta_{\ell-1}) - Z_\ell(\theta_{\ell-1} + s^1_\ell)\big].$$

A MC discretization method of the last equation now has a small variance if $Z_{\ell-1}(\theta_{\ell-1}) - Z_\ell(\theta_{\ell-1} + s^1_\ell)$ and s^2_ℓ is small. This follows, since (8.6) is a linear system similarly to the system (7.35) in Chapter 7. However, we have to potentially use multiple Newton steps, since the expression $o(\|\delta_\ell\|)$ might be large on coarse grids. This clearly complicates the analysis.
This root-finding algorithm can obviously be used for the optimization problem

$$\min_\theta \mathbb{E}[Z(\theta)],$$

where we apply Newton's method to the first order optimality conditions

$$\mathbb{E}\big[\nabla^2_\theta Z(\theta)\big] s = -\mathbb{E}[\nabla_\theta Z(\theta)].$$

Here $\nabla^2_\theta Z$ denotes the Hessian of Z w.r.t. the parameter θ and s the Newton step. The proposed method for stochastic optimization is a second order method and thus differs from the well–known first order stochastic gradient descent methods, where multilevel variance reduction was already incorporated [43, 51, 94, 95, 138].

Conditional value at risk and risk averse optimization problem. The conditional value-at-risk (CVaR) (or average value-at-risk) describes the tail expectation of a random variable and is a coherent risk measure, see [127, Chapter 6] or [110, Chapter 2] for more information on risk measures. The authors of [119] show that computing the CVaR can be formulated as non–smooth minimization problem

$$\mathrm{CVaR}(Z) := \min_t t + \frac{1}{1-\tau}\mathbb{E}[(Z-t)_+], \tag{8.7}$$

where $\tau \in (0,1)$ is the risk level and $(x)_+ := \max\{x, 0\}$. The optimal t^* is called the value-at-risk, whereas the minimum is the CVaR. The authors of [86] developed a method to compute the cumulative distribution function using the MLMC estimator. An approximation to the CVaR is then computed using a post processing step. Giles et al. [58] used a multilevel estimator for a nested expectation involving probabilities and then use a root-finding algorithm to obtain the value-at-risk and the CVaR. A future research direction would be to compute the minimum in (8.7) with a similar approach we outlined in Chapter 7, which of course has to be adapted to deal with the non–linearity of $(\cdot)_+$. We write down the optimality conditions of (8.7) for $t_1^*, \dots, t_L^*$ as follows

$$\mathbb{P}(Z_\ell \leq t_\ell^*) = \tau.$$

Now linearise the left–hand side to introduce an update $\delta_\ell = t_\ell^* - t_{\ell-1}^*$ such that

$$\mathbb{P}(Z_\ell \leq t_\ell^*) = \mathbb{P}(Z_\ell \leq t_{\ell-1}^* + \delta_\ell) = \mathbb{P}(Z_\ell \leq t_{\ell-1}^*) + p_\ell(t_{\ell-1}^*)\delta_\ell + o(\delta_\ell),$$

where p_ℓ is the probability density function of Z_ℓ. The condition for the Newton step s_ℓ^1 then reads

$$p_\ell(t_{\ell-1}^*)s_\ell^1 = \tau - \mathbb{P}(Z_\ell \leq t_{\ell-1}^*) = \mathbb{P}(Z_{\ell-1} \leq t_{\ell-1}^*) - \mathbb{P}(Z_\ell \leq t_{\ell-1}^*).$$

If δ_ℓ is not small enough, we use another Newton step

$$p_\ell(t_{\ell-1}^* + s_\ell^1)s_\ell^2 = \mathbb{P}(Z_{\ell-1} \leq t_{\ell-1}^*) - \mathbb{P}(Z_\ell \leq t_{\ell-1}^* + s_\ell^1).$$

We discretize the left and right–hand side using an MC estimator for the quantile and density independently of $\widehat{t}_{\ell-1}^*$, which we assume was already estimated. This then gives

$$\widehat{p}_\ell(\widehat{t}_{\ell-1}^*)\widehat{s}_\ell = \widehat{\mathbb{P}}(Z_{\ell-1} \leq \widehat{t}_{\ell-1}^*) - \widehat{\mathbb{P}}(Z_\ell \leq \widehat{t}_{\ell-1}^*).$$

If the estimator for the density and the difference of the quantile estimators is sufficiently small, then this should yield a variance reduction comparable to the standard MLMC estimator.
The goal is of course to apply this method to a risk averse optimal control problem, which is given as follows

$$\begin{aligned} \min_u \quad & J(u) = \mathrm{CVaR}(\|y(u) - y_d\|^2) + \frac{\alpha}{2}\|u\|^2 \\ \text{such that } \mathbb{P}\text{–a.s.} \quad & y(u,\omega) = S(\omega)u. \end{aligned}$$

This problem without variance reduction was studied in [85].

Risk neutral optimal control problem with control constraints. The optimal control problem in Chapter 7 does not have control constraints. With control constraints this problem is given as follows

$$\begin{aligned} \min_u \qquad & J(u) = \frac{1}{2}\mathbb{E}\big[\|y(u) - y_d\|^2\big] + \frac{\alpha}{2}\|u\|^2 \\ \text{such that } \mathbb{P}\text{–a.s.} \qquad & y(u,\omega) = S(\omega)u, \\ & a \le u \le b, \end{aligned} \tag{8.8}$$

where $a \le b$ are two real–valued numbers or functions in $L^2(D)$. Now assume that the solution $u_{\ell-1}$, that is if we discretize S by $S_{\ell-1}$, is known. Then for $\delta_\ell := u_\ell - u_{\ell-1}$ the corresponding optimal control problem is

$$\begin{aligned} \min_{\delta_\ell} \qquad & J(\delta_\ell) = \frac{1}{2}\mathbb{E}\big[\|y_\ell(u_{\ell-1} + \delta_\ell) - y_d\|^2\big] + \frac{\alpha}{2}\|u_{\ell-1} + \delta_\ell\|^2 \\ \text{such that } \mathbb{P}\text{–a.s.} \quad & y_\ell(u_{\ell-1} + \delta_\ell, \omega) = S_\ell(\omega)(u_{\ell-1} + \delta_\ell), \\ & a \le u_{\ell-1} + \delta_\ell \le b. \end{aligned}$$

We now rewrite this to obtain

$$\begin{aligned} \min_{\delta_\ell} \qquad & J(\delta_\ell) = \frac{1}{2}\mathbb{E}\big[\|y_\ell(\delta_\ell) - \widetilde{y_d}\|^2\big] + \frac{\alpha}{2}\|u_{\ell-1} + \delta_\ell\|^2 \\ \text{such that } \mathbb{P}\text{–a.s.} \quad & y_\ell(\delta_\ell, \omega) = S_\ell(\omega)\delta_\ell, \\ & a - u_{\ell-1} \le \delta_\ell \le b - u_{\ell-1}, \end{aligned} \tag{8.9}$$

where $\widetilde{y_d} := y_\ell(u_{\ell-1}) - y_d$. The difference compared to the original (8.8) is that the control constraints depend on the previous level, $\widetilde{y_d}$ is now random and the regularization term includes $u_{\ell-1}$. Crucially, if $S_\ell = S_{\ell-1}$, then the unique solution of (8.9) is $\delta_\ell = 0$. Hence for $S_\ell \approx S_{\ell-1}$ the norm of the increment is small and it is reasonable to expect that the variance of the MC estimator for the mean in J in (8.9) is also small. The result is that we should need few samples for a good approximation of δ_ℓ. An approach similar to Chapter 7 could be used to solve this problem. For example, if a version of [73, Theorem 3.5] holds in our setting when replacing u with δ, h^2 with the MC error, squaring and changing the norm then gives

$$\mathbb{E}\Big[\|\delta_\ell - \widehat{\delta}_\ell\|^2\Big] \le c\frac{1}{m_\ell}(\mathbb{E}\big[\|y_\ell(\delta_\ell)\|^2\big] + \|\delta_\ell\|^2) \le c\frac{2^{-2\gamma\ell}}{m_\ell}.$$

It is not clear whether such a bound holds, however, if it does then we have a variance reduction similar to the standard MLMC estimator.

Bibliography

[1] R. J. Adler and J. E. Taylor. *Random fields and Geometry*. Springer New York, 2007.

[2] M. Aigner. *Diskrete Mathematik*. Vieweg+Teubner Verlag, sixth edition, 2006.

[3] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas. Mean-Variance Risk-Averse Optimal Control of Systems Governed by PDEs with Random Parameter Fields Using Quadratic Approximations. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):1166–1192, 2017.

[4] A. A. Ali, E. Ullmann, and M. Hinze. Multilevel Monte Carlo Analysis for Optimal Control of Elliptic PDEs with Random Coefficients. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):466–492, 2017.

[5] M. Asadzadeh, A. H. Schatz, and W. Wendland. A new approach to Richardson extrapolation in the finite element method for second order elliptic problems. *Mathematics of Computation*, 78(268):1951–1973, 2009.

[6] R. B. Ash. *Probability and Measure Theory*. Academic Press, second edition, 2000.

[7] G. Astarita. *Thermodynamics*. Springer, New York, 1989.

[8] P. Bhattacharya and P. Burman. *Theory and Methods of Statistics*. Academic Press, 2016.

[9] C. Bierig and A. Chernov. Convergence analysis of multilevel Monte Carlo variance estimators and application for random obstacle problems. *Numer. Math.*, 130(4):579–613, 2015.

[10] P. Billingsley. *Probability and measure*. Wiley, Hoboken, NJ, anniversary edition, 2012.

[11] H. Blum, Q. Lin, and R. Rannacher. Asymptotic error expansion and Richardson extrapolation for linear finite elements. *Numer. Math.*, 49(1):11–37, 1986.

[12] V. I. Bogachev. *Gaussian Measures*. AMS, Providence, RI, 1998.

[13] J. A. Bondy and U. Murty. *Graph theory with applications*. Macmillan, London, 1976.

[14] A. Borzì and V. Schulz. Multigrid Methods for PDE Optimization. *SIAM Review*, 51(2):361–395, 2009.

[15] A. Borzì and V. Schulz. *Computational Optimization of Systems Governed by Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2011.

[16] A. Borzì, V. Schulz, C. Schillings, and G. von Winckel. On the treatment of distributed uncertainties in PDE-constrained optimization. *GAMM-Mitteilungen*, 33(2):230–246, 2010.

[17] A. Borzì and G. von Winckel. Multigrid Methods and Sparse-Grid Collocation Techniques for Parabolic Optimal Control Problems with Random Coefficients. *SIAM Journal on Scientific Computing*, 31(3):2172–2192, 2009.

[18] A. Borzì and G. von Winckel. A POD framework to determine robust controls in PDE optimization. *Computing and Visualization in Science*, 14(3):91–103, 2011.

[19] S. Boucheron, G. Lugosi, and O. Bousquet. *Concentration Inequalities*, Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures, pages 208–240. Springer, Berlin, Heidelberg, 2004.

[20] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[21] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, third edition, 2008.

[22] C. Brezinski. Convergence Acceleration during the 20th Century. *J. Comput. Appl. Math.*, 122(1–2):1–21, 2000.

[23] R. Bulirsch and J. Stoer. Numerical Treatment of Ordinary Differential Equations by Extrapolation Methods. *Numerische Mathematik*, 8:1–13, 1966.

[24] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.

[25] J. Charrier. Strong and Weak Error Estimates for Elliptic Partial Differential Equations with Random Coefficients. *SIAM Journal on Numerical Analysis*, 50(1):216–246, 2012.

[26] J. Charrier, R. Scheichl, and A. L. Teckentrup. Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods. *SIAM Journal on Numerical Analysis*, 51(1):322–352, 2013.

[27] P. Chen and A. Quarteroni. Weighted Reduced Basis Method for Stochastic Optimal Control Problems with Elliptic PDE Constraint. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):364–396, 2014.

[28] N. Chopin and O. Papaspiliopoulos. *An Introduction to Sequential Monte Carlo*. Springer, Cham, 2020.

[29] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, 2002.

[30] K. Cliffe, I. Graham, R. Scheichl, and L. Stals. Parallel Computation of Flow in Heterogeneous Media Modelled by Mixed Finite Elements. *Journal of Computational Physics*, 164(2):258 – 282, 2000.

[31] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(1):3–15, 2011.

[32] L. W. Cook, J. P. Jarrett, and K. E. Willcox. Generalized information reuse for optimization under uncertainty with non-sample average estimators. *International Journal for Numerical Methods in Engineering*, 115(12):1457–1476, 2018.

[33] N. A. C. Cressie. *Statistics for spatial data*. Wiley, New York, revised edition, 1993.

[34] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, second edition, 2014.

[35] G. Dahlquist and A. Björck. *Numerical Methods in Scientific Computing, Volume I*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008.

[36] M. Dambrine, C. Dapogny, and H. Harbrecht. Shape Optimization for Quadratic Functionals and States with Random Right-Hand Sides. *SIAM Journal on Control and Optimization*, 53(5):3081–3103, 2015.

[37] M. Dashti and A. M. Stuart. *The Bayesian Approach to Inverse Problems*, Handbook of Uncertainty Quantification, pages 311–428. Springer, Cham, 2017.

[38] P. J. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Academic Press, second edition, 1984.

[39] J. De los Reyes. *Numerical PDE-Constrained Optimization*. SpringerBriefs in Optimization. Springer, 2015.

[40] G. de Marsily. *Quantitative hydrogeology*. Academic Press, Orlando, 1986.

[41] A. Defazio, F. Bach, and S. Lacoste-Julien. *SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives*, Advances in Neural Information Processing Systems, volume 27, pages 1646–1654. Curran Associates, Inc., 2014.

[42] F. M. Dekking. *A modern introduction to probability and statistics*. Springer, London, 2005.

[43] S. Dereich and T. Müller-Gronbach. General multilevel adaptations for stochastic approximation algorithms of Robbins–Monro and Polyak–Ruppert type. *Numerische Mathematik*, 142(2):279–328, 2019.

[44] G. Detommaso, T. Dodwell, and R. Scheichl. Continuous Level Monte Carlo and Sample-Adaptive Model Hierarchies. *SIAM/ASA Journal on Uncertainty Quantification*, 7(1):93–116, 2019.

[45] J. L. Devore and K. N. Berk. *Modern Mathematical Statistics with Applications*. Springer, New York, second edition, 2012.

[46] J. Dick, F. Y. Kuo, and I. H. Sloan. High-dimensional integration: The quasi-Monte Carlo way. *Acta Numerica*, 22:133–288, 2013.

[47] C. R. Dietrich and G. N. Newsam. Fast and Exact Simulation of Stationary Gaussian Processes through Circulant Embedding of the Covariance Matrix. *SIAM Journal on Scientific Computing*, 18(4):1088–1107, 1997.

[48] A. Doucet, N. Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer New York, 2001.

[49] S. Engel, D. Hafemeyer, C. Münch, and D. Schaden. An application of sparse measure valued Bayesian inversion to acoustic sound source identification. *Inverse Problems*, 35(7):075005, 2019.

[50] L. C. Evans. *Partial differential equations*. American Mathematical Society, Providence, R.I., second edition, 2010.

[51] N. Frikha. Multi-level stochastic approximation algorithms. *Ann. Appl. Probab.*, 26(2):933–985, 2016.

[52] S. Garreis and M. Ulbrich. Constrained Optimization with Low-Rank Tensors and Applications to Parametric Problems with PDEs. *SIAM Journal on Scientific Computing*, 39(1):A25–A54, 2017.

[53] C. Geiersbach and W. Wollner. A Stochastic Gradient Method With Mesh Refinement for PDE-Constrained Optimization Under Uncertainty. *SIAM Journal on Scientific Computing*, 42(5):A2750–A2772, 2020.

[54] R. Ghanem, D. Higdon, and H. Owhadi, editors. *Handbook of Uncertainty Quantification*. Springer International Publishing, 2017.

[55] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer, New York, 1991.

[56] M. B. Giles. Multi-level Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[57] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 2015.

[58] M. B. Giles and A.-L. Haji-Ali. Multilevel Nested Simulation for Efficient Risk Estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):497–525, 2019.

[59] P. Glasserman. *Monte Carlo methods in financial engineering*. Springer, New York, 2003.

[60] P. W. Glynn and R. Szechtman. Some New Perspectives on the Method of Control Variates. In K.-T. Fang, H. Niederreiter, and F. J. Hickernell, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2000*, pages 27–49, Springer, Berlin, Heidelberg, 2002.

[61] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, fourth edition, 2013.

[62] A. A. Gorodetsky, G. Geraci, M. S. Eldred, and J. D. Jakeman. A generalized approximate control variate framework for multifidelity uncertainty quantification. *Journal of Computational Physics*, 408:109257, 2020.

[63] I. G. Graham, R. Scheichl, and E. Ullmann. Mixed finite element analysis of lognormal diffusion and multilevel Monte Carlo methods. *Stoch. Partial Differ. Equ. Anal. Comput.*, 4(1):41–75, 2016.

[64] J. Groß. *Linear Regression*. Springer Berlin Heidelberg, 2003.

[65] M. Gunzburger and J. Ming. Optimal Control of Stochastic Flow over a Backward-Facing Step Using Reduced-Order Modeling. *SIAM Journal on Scientific Computing*, 33(5):2641–2663, 2011.

[66] P. A. Guth, V. Kaarnioja, F. Y. Kuo, C. Schillings, and I. H. Sloan. A quasi-Monte Carlo Method for an Optimal Control Problem Under Uncertainty, *arXiv*:1910.10022, 2019.

[67] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer, Berlin, 1985.

[68] A.-L. Haji-Ali, F. Nobile, and R. Tempone. Multi-index Monte Carlo: when sparsity meets sampling. *Numerische Mathematik*, 132(4):767–806, 2016.

[69] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009.

[70] S. Heinrich. Multilevel Monte Carlo Methods. In S. Margenov, J. Waśniewski, and P. Yalamov, editors, *Large-Scale Scientific Computing*, pages 58–67, Springer, Berlin, Heidelberg, 2001.

[71] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49:409–436, 1952.

[72] M. Hinze. A Variational Discretization Concept in Control Constrained Optimization: The Linear-Quadratic Case. *Computational Optimization and Applications*, 30(1):45–61, 2005.

[73] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Mathematical modelling–theory and applications. Springer, 2009.

[74] H. Hoel, E. von Schwerin, A. Szepessy, and R. Tempone. Adaptive multilevel monte carlo simulation. In B. Engquist, O. Runborg, and Y.-H. R. Tsai, editors, *Numerical Analysis of Multiscale Computations*, pages 217–234, Springer, Berlin, Heidelberg, 2012.

[75] L. Hou, J. Lee, and H. Manouzi. Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs. *Journal of Mathematical Analysis and Applications*, 384(1):87 – 103, 2011.

[76] J. Humpherys, P. Redd, and J. West. A fresh look at the Kalman filter. *SIAM Rev.*, 54(4):801–823, 2012.

[77] O. Kallenberg. *Foundations of modern probability*. Springer, New York, first edition, 1997.

[78] C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. Society for Industrial and Applied Mathematics, 1995.

[79] A. Klenke. *Probability Theory: A Comprehensive Course*. Springer, London, second edition, 2014.

[80] R. Kornhuber and E. Youett. Adaptive multilevel Monte Carlo methods for stochastic variational inequalities. *SIAM Journal on Numerical Analysis*, 56(4):1987–2007, 2018.

[81] D. P. Kouri. *An Approach for the Adaptive Solution of Optimization Problems Governed by Partial Differential Equations with Uncertain Coefficients*. PhD thesis, Rice University, 2012.

[82] D. P. Kouri. A Multilevel Stochastic Collocation Algorithm for Optimization of PDEs with Uncertain Coefficients. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):55–81, 2014.

[83] D. P. Kouri, M. Heinkenschloss, D. Ridzal, and B. G. van Bloemen Waanders. A Trust-Region Algorithm with Adaptive Stochastic Collocation for PDE Optimization under Uncertainty. *SIAM Journal on Scientific Computing*, 35(4):A1847–A1879, 2013.

[84] D. P. Kouri, M. Heinkenschloss, D. Ridzal, and B. G. van Bloemen Waanders. Inexact Objective Function Evaluations in a Trust-Region Algorithm for PDE-Constrained Optimization under Uncertainty. *SIAM Journal on Scientific Computing*, 36(6):A3011–A3029, 2014.

[85] D. P. Kouri and T. M. Surowiec. Risk-Averse PDE-Constrained Optimization Using the Conditional Value-At-Risk. *SIAM Journal on Optimization*, 26(1):365–396, 2016.

[86] S. Krumscheid and F. Nobile. Multilevel Monte Carlo Approximation of Functions. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1256–1293, 2018.

[87] F. Y. Kuo, R. Scheichl, C. Schwab, I. H. Sloan, and E. Ullmann. Multilevel Quasi-Monte Carlo methods for lognormal diffusion problems. *Math. Comp.*, 86(308):2827–2860, 2017.

[88] S. S. Lavenberg and P. D. Welch. A Perspective on the Use of Control Variables to Increase the Efficiency of Monte Carlo Simulations. *Management Science*, 27(3):322–335, 1981.

[89] V. Lemaire and G. Pagès. Multilevel Richardson-Romberg extrapolation. *Bernoulli*, 20(3):1029–1067, 2017.

[90] C. Lemieux. *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer, New York, NY, 2009.

[91] R. M. Lewis and S. G. Nash. Model Problems for the Multigrid Optimization of Systems Governed by Differential Equations. *SIAM Journal on Scientific Computing*, 26(6):1811–1837, 2005.

[92] G. J. Lord, C. E. Powell, and T. Shardlow. *An Introduction to Computational Stochastic PDEs*. Cambridge University Press, 2014.

[93] M. Martin, S. Krumscheid, and F. Nobile. Analysis of stochastic gradient methods for PDE-constrained optimal control problems with uncertain parameters. Technical report, MATHICSE, Écublens, 2018.

[94] M. Martin, F. Nobile, and P. Tsilifis. A Multilevel Stochastic Gradient method for PDE-constrained Optimal Control Problems with uncertain parameters, *arXiv*:1912.11900, 2019.

[95] M. C. Martin and F. Nobile. PDE-constrained optimal control problems with uncertain parameters using SAGA, *arXiv*:1810.13378, 2020.

[96] N. Matloff. *Statistical regression and classification: from linear models to machine learning*. CRC Press, 2017.

[97] C. Mbaye, G. Pagès, and F. Vrins. An antithetic approach of multilevel Richardson-Romberg extrapolation estimator for multidimensional SDEs. In *Numerical analysis and its applications*, volume 10187 of *Lecture Notes in Comput. Sci.*, pages 482–491. Springer, Cham, 2017.

[98] G. Meurant. *The Lanczos and Conjugate Gradient Algorithms*. Society for Industrial and Applied Mathematics, 2006.

[99] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of Machine Learning*. The MIT Press, second edition, 2018.

[100] S. G. Nash. A multigrid approach to discretized optimization problems. *Optimization Methods and Software*, 14(1-2):99–116, 2000.

[101] L. W. T. Ng and K. E. Willcox. Monte Carlo Information-Reuse Approach to Aircraft Conceptual Design Optimization Under Uncertainty. *Journal of Aircraft*, 53(2):427–438, 2016.

[102] F. Nobile, R. Tempone, and C. G. Webster. A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[103] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, NY, USA, second edition, 2006.

[104] G. Pagès. Multi-step Richardson-Romberg Extrapolation: Remarks on Variance Control and Complexity. *Monte Carlo Methods and Applications*, 13(1):37 – 70, 2007.

[105] B. Peherstorfer. Multifidelity Monte Carlo Estimation with Adaptive Low-Fidelity Models. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):579–603, 2019.

[106] B. Peherstorfer, M. Gunzburger, and K. Willcox. Convergence analysis of multifidelity Monte Carlo estimation. *Numerische Mathematik*, 139(3):683–707, 2018.

[107] B. Peherstorfer, K. Willcox, and M. Gunzburger. Optimal Model Management for Multifidelity Monte Carlo Estimation. *SIAM Journal on Scientific Computing*, 38(5):A3163–A3194, 2016.

[108] B. Peherstorfer, K. Willcox, and M. Gunzburger. Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization. *SIAM Review*, 60(3):550–591, 2018.

[109] K. B. Petersen and M. S. Pedersen. The Matrix Cookbook, 2012. Technical University of Denmark, Version 20121115, url: http://www2.imm.dtu.dk/pubdb/edoc/imm3274.pdf, Last access 23.12.2020.

[110] G. C. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk*. World Scientific, 2007.

[111] J. Potthoff. Sample properties of random fields III: differentiability. *Communications on Stochastic Analysis*, 4(3):335–353, 2010.

[112] E. Qian, B. Peherstorfer, D. O'Malley, V. V. Vesselinov, and K. Willcox. Multifidelity Monte Carlo estimation of variance and sensitivity indices. *SIAM/ASA J. Uncertain. Quantif.*, 6(2):683–706, 2018.

[113] R. Rannacher. *Extrapolation techniques in the finite element method: A survey*, number MATC7 in Proceedings of the Summer School on Numerical Analysis, Helsinki, Univ. of Tech., 1988.

[114] C. R. Rao, H. Toutenburg, Shalabh, and C. Heumann. *Linear Models and Generalizations*. Springer Series in Statistics. Springer, Berlin, third extended edition, 2008.

[115] C. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, USA, 2006.

[116] J. Rawlings, S. Pantula, and D. Dickey. *Applied regression analysis: a research tool*. Springer texts in statistics. Springer, New York, NY, second edition, 1998.

[117] L. F. Richardson. The Approximate Arithmetical Solution by Finite Differences of Physical Problems Involving Differential Equations, with an Application to the Stresses in a Masonry Dam. *Philosophical Transactions of the Royal Society of London. Series A*, 210:307–357, 1911.

[118] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, New York, 2004.

[119] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[120] W. Romberg. Vereinfachte numerische Integration. *Det Kongelige Norske Videnskabers Selskab Forhandlinger*, 28(7):30–36, 1955.

[121] E. Rosseel and G. N. Wells. Optimal control with stochastic PDE constraints and uncertain controls. *Computer Methods in Applied Mechanics and Engineering*, 213-216:152 – 167, 2012.

[122] G. G. Roussas. *An Introduction to Probability and Statistical Inference*. Academic Press, Boston, second edition, 2015.

[123] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*. Wiley, Hoboken, N.J., second edition, 2008.

[124] A. K. Saibaba, J. Lee, and P. K. Kitanidis. Randomized algorithms for generalized Hermitian eigenvalue problems with application to computing Karhunen–Loève expansion. *Numerical Linear Algebra with Applications*, 23(2):314–339, 2016.

[125] D. Schaden and E. Ullmann. Asymptotic analysis of multilevel best linear unbiased estimators, *arXiv*:2012.03658, 2020. *submitted*.

[126] D. Schaden and E. Ullmann. On Multilevel Best Linear Unbiased Estimators. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):601–635, 2020.

[127] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming*. Society for Industrial and Applied Mathematics, 2009.

[128] M. L. Stein. *Interpolation of Spatial Data*. Springer Series in Statistics. Springer, New York, 1999.

[129] A. M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

[130] T. Sullivan. *Introduction to Uncertainty Quantification*. Springer International Publishing, 2015.

[131] W. Sun and Y.-X. Yuan. *Optimization Theory and Methods*. Springer, New York, NY, 2006.

[132] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA, 2018.

[133] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):483–509, 1990.

[134] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann. Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numerische Mathematik*, 125:569–600, 2013.

[135] H. Tiesler, R. M. Kirby, D. Xiu, and T. Preusser. Stochastic Collocation for Optimal Control Problems with Stochastic PDE Constraints. *SIAM Journal on Control and Optimization*, 50(5):2659–2682, 2012.

[136] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg+Teubner, Wiesbaden, second edition, 2009.

[137] U. Trottenberg, C. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, 2001.

[138] A. Van Barel and S. Vandewalle. Robust Optimization of PDEs with Random Coefficients Using a Multilevel Monte Carlo Method. *SIAM/ASA Journal on Uncertainty Quantification*, 7(1):174–202, 2019.

[139] A. Van Barel and S. Vandewalle. MG/OPT and MLMC for Robust Optimization of PDEs, *arXiv*:2006.01231, 2020.

[140] H. Wackernagel. *Multivariate Geostatistics*. Springer, Berlin, third edition, 2003.

[141] H. Wang and M. Anderson. *Introduction to Groundwater Modeling: Finite Difference and Finite Element Methods*. Academic Press, San Diego, 1982.

[142] S. Weisberg. *Applied Linear Regression*. John Wiley, fourth edition, 2013.

List of Figures

1.1 Chapters of this thesis . . . 17
2.1 Gaussian random fields with mean zero and Whittle–Matérn covariance in $1d$ . . . 32
2.2 Gaussian random fields with mean zero and Whittle–Matérn covariance in $2d$ . . . 33
2.3 Solutions of the elliptic PDE for different diffusion coefficients in $2d$ . . . 38
4.1 Model groups of different linear unbiased estimators . . . 100
4.2 $\mathbb{V}^{\min}$ of the monomial example . . . 102
4.3 Variance of different estimator for the monomial example . . . 103
4.4 Coefficients of different estimators for the monomial example . . . 104
4.5 $\mathbb{V}^{\min}$ of the noisy monomial example . . . 106
4.6 Variance of different estimator for the noisy monomial example . . . 106
4.7 Coefficients of different estimators for the noisy monomial example . . . 108
5.1 Sections of Chapter 5 described in terms of minimization problems . . . 110
6.1 Convergence of the SAOB to the RE estimator . . . 163
6.2 Coefficients of the SAOB 2 and the MLMC estimator . . . 163
6.3 Coefficients of the SAOB 3 and the RE 3 estimator . . . 164
6.4 Variance and errors of estimators with specific linear combinations . . . 166
6.5 Biases, variances and costs of the PDE example . . . 167
6.6 Asymptotic complexity of different estimators for the PDE example with true cost . . . 169
6.7 Number of high fidelity evaluations for the PDE example with true cost . . . 170
6.8 Asymptotic complexity of different estimators for the PDE example with artificial cost . . . 171
6.9 Number of high fidelity evaluations for the PDE example with artificial cost 171
6.10 Asymptotic complexity of the CV estimators for the PDE example with artificial cost . . . 172
6.11 Coefficients of different estimators for the PDE example with artificial costs 173
7.1 Bias, variance, variance difference and costs of the control . . . 202
7.2 Computed cost versus idealized mean square error of the MC and MLC estimator . . . 204
7.3 Computed cost versus mean square error of the MC and MLC estimator . 206
7.4 True cost versus mean square error of the MC and MLC estimator . . . 206

List of Tables

3.1 Asymptotic cost for MC for different dimensions d for the PDE QoI 47
3.2 Asymptotic cost for MFMC for different dimensions d for the PDE QoI . . 61
4.1 Table comparing the conditions that linear unbiased estimators are BLUE 100
4.2 Mean, variance and correlation for the monomial example 101
4.3 Mean, variance and correlation for the noisy monomial example 105
6.1 Cost of different estimators for the PDE example with true cost 168
6.2 Cost of different estimators for the PDE example with artificial cost 170
7.1 Sample allocation of the MC and MLC estimators 204
7.2 Quotient of the true MSE divided by the idealized MSE 205

List of Symbols

1 Indicator function 75, 97, 136, 137

∇ Gradient 32–34, 36, 37, 39, 167, 176, 201, 212

a Random diffusion field 28–40, 42, 167, 201

B The operator of the optimality conditions 178–181, 184, 190, 192–195

$\widehat{B}$ The MC estimator of the operator of the optimality conditions 182–186, 188, 190, 192–196

C Covariance matrix 24, 25, 48–52, 54, 56, 63–66, 76–79, 81–93, 95, 96, 98, 101, 109, 110, 112–120, 122–126, 128, 129, 131–137, 140–142, 144, 145, 151, 153, 158–162, 164–167, 201, 202, 207–212

$\mathcal{C}$ Covariance operator 28–30, 35, 38, 201

$\mathbb{C}\mathbf{orr}$ Correlation 24, 25

$\mathbb{C}\mathbf{ov}$ Covariance 24–26, 28–30, 35, 45, 52, 54, 57, 59, 81–85, 87, 88, 201, 211

d Physical dimension 30, 32–35, 38, 47, 61, 72, 225, 226

D Equal to $\begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix}$, shifts down a vector 147–149, 151, 153–155, 172, 173

D Physical domain 29–42, 47, 61, 72, 176, 201, 214, 227, 231

D_{obs} Observation domain, an open subset of the physical domain D 34, 38, 42, 167

div Divergence operator applied to a vectorfield 32, 33, 167, 201

$\mathbb{E}$ Expectation of a random variable 23–30, 37–39, 41–46, 48, 67, 72–74, 76, 79, 83, 84, 88, 140, 145–148, 151, 153, 160, 161, 176–188, 190–201, 203, 212–214

f Right–hand side of the optimality conditions 178–181, 183, 184, 190

$\widehat{f}$ MC estimated right–hand side of the optimality conditions 182–184, 194

Δf Difference of two levels of the right–hand side of the optimality conditions 190, 192, 194–196

$\Delta\widehat{f}$ MC estimated difference of two levels of the right–hand side of the optimality conditions 190, 192, 194, 196

$\mathcal{F}$ Sigma–algebra 19, 20, 22

g Boundary function on ∂D 33, 34

$\mathcal{G}$ Sigma–algebra on H 20, 21

h Mesh size of a triangulation $\mathcal{T}$ 36–39, 42, 47

H Hilbert space 20–32, 88, 89, 227

H_0^1 Sobolev space with first weak derivative and zero value at the boundary 34–37, 40, 41, 176

J_δ Variance of the relaxed BLUE with covariance matrix Ψ_δ 115–121

k Whittle–Matérn Covariance kernel 29–31, 35

K K parameter of the ACV–KL estimator 64–67, 97, 98, 103

K Number of model groups, $K := 2^L - 1$ 74–76, 80–82, 109–130, 132, 134–136, 140–144, 158, 161, 162, 209, 211, 212

ℓ Correlation length of a Whittle–Matérn covariance kernel 29, 31–33, 38, 167, 201

L Lebesgue space L^p 22–28, 30, 31, 34–39, 41–43, 46, 88, 176, 191, 201, 214

m Number of samples 21, 45–55, 59–72, 74–99, 102, 109–126, 128, 129, 133–137, 140, 142–144, 150–153, 156, 157, 166, 182–200, 202, 203, 209, 210, 214

n Number of CG–iterations 185, 186, 188, 200

N Truncation index of the Karhunen–Loève expansion 30–33, 42

N L parameter of the ACV–KL estimator 64–67, 97, 98, 103

P Prolongation operator for vectors with zero extension 74–79, 81, 83–86, 89, 109, 110, 113–115, 117–120, 124–134, 136, 140, 141, 144, 209–211

$\mathbb{P}$ Probability measure 19–23, 25–27, 32, 34, 37, 39, 41, 42, 44, 45, 135–137, 146, 157, 167, 176, 178–181, 191, 213, 214

q Generic random bound 178–187, 191–193

q_{mean} Number of terms of the RE expansion for the mean 146–148, 151–157, 160, 161, 167

q_{path} Number of terms of the pathwise RE expansion 146–148, 151–157, 159–162, 167, 168

Q Equal to $\mathbb{E}[S^*S]$, where S is the control to state map and S^* its adjoint 178, 180, 181, 183, 184, 190, 192, 193, 195, 196

$\widehat{Q}$ MC estimator of Q 182–184, 186, 192–196

ΔQ Difference of Q on two consecutive discretization levels 190, 192–196

$\Delta\widehat{Q}$ MC estimator of ΔQ 190, 192–196

R Restriction operator, removes entries of a vector 74, 76, 78–81, 83–85, 89, 93, 96, 109, 113, 115, 117–120, 133, 134

S Model group, a non–empty subset of $\{1, \dots, L\}$ 74–78, 80–100, 102, 103, 106, 107, 110–112, 114–120, 122–126, 129, 131, 136, 137, 140, 142–145, 150, 158, 161, 172, 173, 209–212

S The control to state map of the optimal control problem 176, 178–193, 198–203, 213, 214

S^* Adjoint of S 178–193, 199–203

u Right–hand side of a PDE and control 33–35, 37, 38, 176–184, 187–203, 213, 214

$\widehat{u}^{\mathrm{MC}}$ MC estimator of the optimal control u 182–189, 191, 200, 202–205

$\widehat{u}^{\mathrm{MC,\,CG}}$ MC estimator of the optimal control u computed with the CG–method 185–188

$\widehat{u}^{\mathrm{MLC}}$ MLC estimator of the optimal control u 190–200, 204, 205

$\widehat{u}^{\mathrm{MLC,\,CG}}$ MLC estimator of the optimal control u computed with the CG–method 198–200

U_Z Set of used models 75, 76, 78–80, 83, 85, 114, 126

U_α Set of models that are required for the evaluation 75, 80, 111, 114, 126

U_S Set of model groups that are evaluated 75–79, 83–86, 88–90, 109, 115, 119–122

U_β Set of model groups where the coefficient of the linear combination is not zero 122–124, 126–135

v RE vector, linearly combines models to achieve an improved asymptotic accuracy 147–157, 159–161, 167–173

$\mathbb{V}$ Variance 23–28, 38, 39, 42–52, 54–56, 58–60, 62–65, 67–72, 76, 77, 79, 85–91, 93, 94, 98, 99, 109, 111, 112, 114, 122, 123, 140, 141, 144, 145, 151, 159, 161, 162, 165, 166, 183, 184, 191–193, 198, 202, 203, 208, 209, 211, 212

$\mathbb{V}^{\mathrm{min}}$ Lower bound for the variance 51, 52, 61, 62, 67, 71, 73, 85–89, 94, 98, 101–103, 105, 106, 112, 113, 129, 139, 157–161, 165, 225

V^{FE} Linear finite element space 37, 39, 40, 231

w Cost or work of a model 21, 42, 47, 49, 50, 55–62, 65, 67–71, 111, 113–115, 121, 122, 126, 128, 129, 136, 137, 145, 151, 152, 156–160, 162, 169, 201

W Cost or work of a model group 110–113, 115, 119–125, 128, 131–136, 140–142, 144, 151, 158, 161, 164, 208–212

$\mathbb{W}$ Cost or work 21, 46–50, 56, 59–61, 68–70, 111, 112, 114, 115, 119, 123, 124, 133–136, 140–145, 152, 156–161, 164, 168, 187–189, 199–203

$\mathbb{W}^{\mathrm{budget}}$ Computational budget 55, 56, 58–60, 62, 65, 68, 69, 71, 110–115, 119, 121–126, 129, 133–137, 162

y Solution of the elliptic PDE 33–40, 42, 167, 176, 178, 179, 181, 182, 189, 191, 201, 213, 214

y_d Desired state of the optimal control problem 176, 178–184, 189–198, 200, 201, 213, 214

Y Generic random variable 21–27, 48, 88, 93, 94

Z Quantity of interest 20–27, 32, 34, 38, 39, 41–55, 57–78, 80–84, 87–107, 111, 113–115, 121, 123, 128, 129, 136, 137, 139–141, 144–148, 150, 151, 153, 157–162, 164–167, 173, 191, 198, 207–213, 230, 231

α Bias parameter of a linear estimator 74–80, 85–87, 109–122, 124–129, 131–134, 136, 139–145, 153, 155, 156, 158–162, 168–170, 172, 208–210

α Tikhonov regularization parameter of the optimal control problem 176, 178, 179, 181–187, 189–192, 194, 196, 197, 199, 201, 203, 213, 214

β Coefficients of a linear combination 23–26, 39, 40, 48–50, 52–54, 62–67, 71, 74–80, 84–86, 88–97, 99, 103, 105, 107, 109, 110, 114, 115, 118–137, 139–142, 144, 145, 150, 158, 159, 161–166, 172, 174, 208–212

γ Generic rate 146–157, 159–162, 164–168, 172, 173, 180, 181, 185–188, 190–200, 202, 203, 214

γ_{Bias} Bias rate 46, 47, 58–61, 70, 140–142, 144–146, 151, 152, 168, 170, 200, 207, 208

γ_{Cost} Cost rate 46, 47, 58–61, 70, 71, 141, 142, 144, 145, 151, 152, 156, 157, 160, 161, 167–173, 187, 188, 199–202, 207, 208, 211

γ_{Var} Variance rate 58–61, 70–72, 141, 142, 144–146, 151, 152, 164, 165, 168, 170, 200, 211

δ Difference of optimal controls on two consecutive discretizations 190–198, 214

$\widehat{\delta}$ MC estimator for δ 190, 191, 194–200, 214

$\widehat{\delta}^{\text{CG}}$ MC estimator for δ computed with the CG–method 198–200

η Random noise 80–84, 112

λ Eigenvalue of the Karhunen–Loève expansion 28–31, 42

μ Mean parameter, $\mu_\ell := \mathbb{E}[Z_\ell]$ for all $\ell \in \{1, \dots, L\}$ 23, 39, 48–54, 61–64, 66, 67, 71, 73–76, 78–85, 87, 88, 91, 92, 95, 101, 104, 105, 108, 114, 115, 140, 145, 151, 153, 160, 161, 209, 210, 212, 230

$\widehat{\mu}$ Estimator for the mean μ 43–45, 51, 72, 74–77, 79, 82–84, 87, 90, 91, 100, 103, 111, 112, 114, 115, 122, 140–144, 158, 159, 161, 166, 168

$\widehat{\mu}^{\text{ACV–IS}}$ Approximate Control Variate Independent Sample estimator 63, 65–67, 94, 95

$\widehat{\mu}^{\text{ACV–KL}}$ Approximate Control Variate KL estimator 64–67, 97, 98

$\widehat{\mu}^{\text{ACV–MF}}$ Approximate Control Variate Multifidelity estimator 63–67, 95, 96

$\widehat{\mu}^{\text{B}}$ Best linear unbiased estimator 78, 79, 82–92, 94–99, 109, 111, 112, 114

$\widehat{\mu}^{\text{CV}}$ Control Variate estimator 48–53, 62, 66, 86, 91

$\widehat{\mu}^{\text{FC}}$ Full coupling estimator 99

$\widehat{\mu}^{\text{MC}}$ Monte Carlo estimator 45–47, 49–51, 56, 61, 90, 111, 122, 144, 168, 170

$\widehat{\mu}^{\text{MFMC}}$ Multifidelity Monte Carlo estimator 53–56, 58–62, 92–94, 96, 122, 128, 168, 170

$\widehat{\mu}^{\text{MLMC}}$ Multilevel Monte Carlo estimator 67–71, 94, 98, 99, 145, 168, 170

$\widehat{\mu}^{\text{RE}}$ Richardson extrapolation estimator 150, 152, 153, 156, 157, 162, 168, 170, 172

$\widehat{\mu}^{\text{SAOB}}$ Sample allocation optimal BLUE 111, 112, 129, 143–145, 152, 157, 160–162, 164, 168–171

ν Smoothness parameter of the Whittle–Matérn covariance kernel 29, 31–33, 35, 38, 167, 201

ξ Random variable of the Karhunen–Loève expansion 28–31, 42

ρ Entry of the correlation matrix P 24, 48, 49, 53–59, 61, 62, 92–94, 96, 98, 99, 101, 105, 122, 128

P Correlation matrix 24, 93, 101, 105

σ^2 Variance of one model in $\{Z_1, \ldots, Z_L\}$ 23, 47–52, 54–59, 61–64, 80, 87–90, 92, 93, 96, 101, 105, 115, 122, 128

σ^2 Variance parameter of the Whittle–Matérn covariance kernel 29, 31, 32, 167, 201

$\mathcal{T}$ Triangulation of the domain D 36, 37, 227

φ Nodal basis function of the finite element space V^{FE} 39, 40

ψ Basis function of the Karhunen–Loève expansion 28–31, 42

Ψ System matrix for the BLUE 78, 79, 82–85, 93, 96, 109–113, 115–118, 120, 124, 133, 134

Ψ_δ System matrix for the BLUE plus δ times the identity matrix I 115–120

ω Elementary event 20–23, 29, 34, 35, 37–42, 81, 101, 105, 146, 147, 157, 159, 160, 162, 165, 167, 168, 170, 176, 178–187, 189–192, 201, 213, 214

Ω Set of elementary events 19–25, 27, 32, 34, 36, 37, 39, 41, 42, 176, 212

List of Abbreviations

a.s. almost surely 22, 181, 191, 213, 214

ACV Approximate Control Variate 65–67, 73, 85, 90, 111, 142, 169, 173, 174

ACV–IS Approximate Control Variate Independent Sample 63, 65, 66, 94, 95, 100, 103

ACV–KL Approximate Control Variate KL 64–67, 97–100, 103, 173, 228

ACV–MF Approximate Control Variate Multifidelity 63–66, 95–100, 103, 173

BLUE Best linear unbiased estimator 15–17, 73, 78–80, 82–101, 103, 105, 107, 109–113, 116, 121, 122, 207–209, 211, 226, 228, 230, 231, 233

CG Conjugate gradient 185, 187–189, 198, 203, 204, 228–230

CV Control Variate 48–53, 56, 61, 62, 64–67, 71, 73, 85, 88, 90, 91, 99, 100, 103, 104, 106, 108, 139, 225

FC Full coupling 99, 100, 103, 105, 107

i.i.d. Independent identically distributed 21, 28, 45, 51, 74, 80, 81, 111, 112, 132, 134, 135, 137, 142, 143, 162, 182

KKT Karush-Kuhn-Tucker 68, 69, 77, 79, 91, 114, 119, 123, 158

KLE Karhunen–Loève Expansion 28–33, 39, 42

MC Monte Carlo 45–52, 56, 60, 61, 67, 70, 71, 77, 88–90, 92, 95, 97, 99, 100, 103, 105, 106, 109, 111, 113, 119, 121, 122, 128, 129, 132, 133, 136, 137, 139, 142, 144, 145, 168, 169, 175, 177, 182–184, 187–193, 198, 200–207, 210, 212–214, 225–230

MFMC Multifidelity Monte Carlo 53–56, 58–63, 66, 69–71, 90, 92–94, 96, 97, 99, 100, 102, 103, 105, 111, 121, 122, 128, 129, 132, 133, 136, 137, 139, 142, 168, 169, 172–174, 211, 226

MLC Multilevel Monte Carlo for the control 175, 190, 191, 196, 198–207, 225, 226, 229

MLMC Multilevel Monte Carlo 67–72, 90, 94, 98–100, 102, 103, 105, 109, 111, 123, 137, 139–142, 144–146, 150, 151, 162, 163, 169, 173–175, 177, 189–191, 198, 202, 207–210, 213, 214, 225

MSE Mean square error 43–49, 59–61, 70, 139–145, 152, 156, 157, 160, 168, 193, 202–206, 226

PDE Partial differential equation 17, 33–35, 41, 42, 47, 61, 67, 68, 72, 139, 146, 167, 169, 175, 176, 180, 184, 202, 207, 225, 226, 229

QoI Quantity of interest 32, 34, 38, 42, 46, 47, 59, 61, 72–74, 80, 94, 101, 105, 114, 139, 144, 159, 160, 162, 165–169, 208, 209, 212, 226

RE Richardson extrapolation 139, 146–153, 155–157, 159, 160, 162–169, 172–174, 209, 211, 225, 228, 229

RMSE Root mean square error 46, 47, 61, 169–173, 188

SAOB Sample allocation optimal BLUE 15–17, 110–112, 121, 126, 128, 129, 132, 139, 141–145, 152, 156, 157, 160, 162–165, 168–171, 173, 174, 207–212, 225

SAOB κ Sample allocation optimal BLUE with coupling κ 112, 143–145, 152, 157, 162, 163, 168, 169, 174