Bayesian Optimal Experimental Design Using Multilevel Monte Carlo

Thesis by
Chaouki ben Issaid

In Partial Fulfillment of the Requirements
For the Degree of
Masters of Science

King Abdullah University of Science and Technology, Thuwal,
Kingdom of Saudi Arabia

May, 2015
The thesis of Chaouki ben Issaid is approved by the examination committee

Committee Chairperson: Prof. Raul Tempone
Committee Member: Prof. Ying Sun
Committee Member: Prof. Fabrizio Bisetti
Committee Member: Prof. Marco Scavino
Committee Member: Dr. Quan Long
ABSTRACT

Bayesian Optimal Experimental Design Using Multilevel Monte Carlo

Chaouki ben Issaid

Experimental design can be vital when experiments are resource-exhaustive and time-consuming. In this work, we carry out experimental design in the Bayesian framework. To measure the amount of information that can be extracted from the data in an experiment, we use the expected information gain as the utility function, which specifically is the expected logarithmic ratio between the posterior and prior distributions. Optimizing this utility function enables us to design experiments that yield the most informative data about the model parameters. One of the major difficulties in evaluating the expected information gain is that it naturally involves nested integration over a possibly high dimensional domain. We use the Multilevel Monte Carlo (MLMC) method to accelerate the computation of the nested high dimensional integral. The advantages are twofold. First, MLMC can significantly reduce the cost of the nested integral for a given tolerance, by using an optimal sample distribution among different sample averages of the inner integrals. Second, the MLMC method imposes fewer assumptions, such as the asymptotic concentration of posterior measures, required for instance by the Laplace approximation (LA). We test the MLMC method using two numerical examples. The first example is the design of sensor deployment for a Darcy flow problem.
governed by a one-dimensional Poisson equation. We place the sensors in the locations where the pressure is measured, and we model the conductivity field as a piecewise constant random vector with two parameters. The second one is chemical Enhanced Oil Recovery (EOR) core flooding experiment assuming homogeneous permeability. We measure the cumulative oil recovery, from a horizontal core flooded by water, surfactant and polymer, for different injection rates. The model parameters consist of the endpoint relative permeabilities, the residual saturations and the relative permeability exponents for the three phases: water, oil and microemulsions. We also compare the performance of the MLMC to the LA and the direct Double Loop Monte Carlo (DLMC). In fact, we show that, in the case of the aforementioned examples, MLMC combined with LA turns to be the best method in terms of computational cost.

**Key words:** Bayesian experimental design, expected information gain, Multilevel Monte Carlo, Laplace approximation, sensor deployment, enhanced oil recovery, core flooding.
ACKNOWLEDGEMENTS

Within the framework of this Master Thesis, I would like to express my deepest thanks to all those who contributed, in their way, to make this thesis so interesting and enriching. I express my sincere gratitude to my supervisor Prof. Raul Tempone for his availability, his advice and his guidance throughout my thesis.

My gratitude also goes to my co-supervisor, Dr Quan Long, for his reviews, suggestions and comments. I am thankful for his cheerfulness and cooperation throughout my thesis.

Last but not least, I would like to thank my family and my dear friends for all their encouragement and support.
# TABLE OF CONTENTS

Examination Committee Approval ........................................... 2
Copyright ................................................................................. 3
Abstract ................................................................................ 4
Acknowledgements ..................................................................... 6
List of Abbreviations ................................................................. 9
List of Symbols ........................................................................ 11
List of Figures .......................................................................... 12
List of Tables ........................................................................... 13

1 Introduction ........................................................................... 14
1.1 Project Framework ............................................................... 14
1.2 Objectives ........................................................................... 16
1.3 Related Work ....................................................................... 16
1.4 Contribution ......................................................................... 17
1.5 Thesis structure .................................................................... 17

2 Bayesian Optimal Experimental Design ................................ 19
2.1 Background ......................................................................... 19
2.2 Framework ......................................................................... 22
2.3 Shannon information gain .................................................... 25
2.4 Examples ............................................................................ 26
2.4.1 Example 1: A linear design problem ................................. 27
2.4.2 Example 2: A one-dimensional nonlinear design problem .. 27
2.4.3 Example 3: Another nonlinear design problem ................. 27
### 3 Laplace approximation

3.1 Laplace approximation in Bayesian inference .............................. 30
3.2 Gaussian approximation of the posterior pdf ............................. 32
3.3 Laplace approximation for the expected information gain estimation . 34

### 4 Multilevel Monte Carlo

4.1 Standard Monte Carlo ............................................................. 37
4.1.1 Monte Carlo for numerical integration ................................. 37
4.1.2 Monte Carlo for the expected information gain estimation ....... 39
4.1.3 Double Loop Monte Carlo .................................................... 40
4.2 Multilevel Monte Carlo ............................................................ 42
4.2.1 Multilevel Monte Carlo for numerical integration ................. 42
4.2.2 Multilevel Monte Carlo for the expected information gain estimation ....................................................... 48
4.2.3 LA combined with MLMC ....................................................... 50

### 5 Numerical Results

5.1 Example 1 ................................................................. 54
5.1.1 Problem set-up .......................................................... 54
5.1.2 Results and discussions .................................................. 58
5.2 Example 2 ................................................................. 63
5.2.1 Notations ............................................................... 63
5.2.2 Problem set-up .......................................................... 64
5.2.3 Mathematical formulation of the forward problem .............. 65
5.2.4 Results and discussion .................................................. 69

### 6 Conclusion

6.1 Summary ................................................................. 74
6.2 Future Work .............................................................. 74

### References

76

### Appendices

82
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{KL}$</td>
<td>Kullback Liebler divergence</td>
</tr>
<tr>
<td>CLT</td>
<td>Central Limit Theorem</td>
</tr>
<tr>
<td>DLMC</td>
<td>Double Loop Monte Carlo</td>
</tr>
<tr>
<td>EOR</td>
<td>Enhanced Oil Recovery</td>
</tr>
<tr>
<td>FEM</td>
<td>Finite Element Method</td>
</tr>
<tr>
<td>i.i.d</td>
<td>independent and identically distributed</td>
</tr>
<tr>
<td>LA</td>
<td>Laplace Approximation</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>MLMC</td>
<td>Multilevel Monte Carlo</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean-square-error</td>
</tr>
<tr>
<td>OOIP</td>
<td>Original Oil In Place</td>
</tr>
<tr>
<td>PDE</td>
<td>Partial Differential Equation</td>
</tr>
<tr>
<td>pdf</td>
<td>probability density function</td>
</tr>
<tr>
<td>RV</td>
<td>Random Variable</td>
</tr>
<tr>
<td>SPDE</td>
<td>Stochastic Partial Differential Equation</td>
</tr>
</tbody>
</table>
UTCHEM  University of Texas CHEMical compositional simulator
**LIST OF SYMBOLS**

- $d$: dimension of parameter vector $\theta$
- $\epsilon$: Gaussian noise
- $\varepsilon_M$: statistical error
- $E_S$: sum of the residuals
- $h(\theta)$: logarithm of the posterior probability density function
- $L$: number of MLMC levels
- $l$: level index
- $N_e$: number of repetitive experiments
- $\mathcal{O}()$: big O notation
- $\Omega$: sample space
- $\mathcal{O}_P()$: big O in probability
- $r_i$: $i^{th}$ residual vector
- $\Sigma$: covariance matrix
- $\Theta$: prior space
- $\theta_t$: true parameter vector
- $TOL$: tolerance
- $tr$: trace of a matrix
- $\xi$: vector of design parameters: the experimental set-up
- $y$: set of observed data points
LIST OF FIGURES

2.1 Posterior pdf of $\theta$ for different values of $N_e$ ............................. 28
4.1 Computational error as a function of the number of samples $N$ .... 39
5.1 $\log_2(\text{variance})$ as a function of the level $l$ ................................. 59
5.2 $\log_2(|\text{mean}|)$ as a function of the level $l$ ................................. 59
5.3 Number of samples per level for different tolerances using MLMC . . 60
5.4 Number of samples per level for different tolerances using MLMC+LA 61
5.5 Cost comparison between the different methods ............................... 61
5.6 Schematic view of the porous media ............................................. 69
5.7 $\log_2(\text{variance})$ as a function of the level $l$ ................................. 70
5.8 $\log_2|\text{mean}|$ as a function of the level $l$ ................................. 70
5.9 Number of samples per level for different tolerances using MLMC . . 71
5.10 Number of samples per level for different tolerances using MLMC+LA 72
5.11 Cost comparison between the different methods ............................... 72
B.1 Basis functions for FEM using $P_1$ elements ................................. 85
LIST OF TABLES

2.1 Optimality criteria for the least squares case ........................................ 21
5.1 Speed-up of the MLMC+LA compared to the other methods .................. 62
5.2 Computational errors of the MLMC+LA and the MLMC methods .......... 62
5.3 Distribution bounds for relative permeability parameters .................... 67
5.4 Summary of the simulation parameters .................................................. 68
5.5 Injection strategy: slugs’ composition details ........................................ 68
5.6 Speed-up of the MLMC+LA compared to the other methods .................. 73
Chapter 1

Introduction

1.1 Project Framework

Most numerical models have parameters such as the speed of light or the Boltzmann constant. Many of these parameters have been experimentally determined and thus often carry a substantial amount of uncertainty. The accuracy of these parameters directly affects how well the models can describe the physical reality.

A possible way to improve these models is to conduct parameter inference. For this purpose, experimental data need to be collected. Since most experiments are expensive, time-consuming and delicate to carry out, it is therefore essential to design experiments that enable us to gain valuable information.

An experiment is usually conducted to obtain some information about a quantity of interest. Gathering this information can be done through elementary experiments known as trials. Another objective of conducting experiments is to understand cause-and-effect relationships within a system. To do so, the operator usually tends to modify the input variables of the system and notice the changes in the system outputs. In other terms, one has to carry out experiments on the system.

In experimental design, the task is how to select the design support points or even better to determine the size $N$ of the design in order to collect enough information about the quantity of interest.
Designing a good experiment is essential because the conclusions that can be deduced from it rely profoundly on the way the data were gathered. The design should prevent systematic error, be accurate and enable error estimation.

Experimental design can be a very efficient tool in engineering applications in order to enhance the product manufacturing process. The application of experimental design procedures early in the process development can lead to improving the process outputs, reduce variability, closer agreement to nominal or target requirements and shorten development time as well as decreasing the overall costs.

Experiments can be designed and performed with different goals. Designing an experiment is highly related to its purpose. For instance, an experiment can be performed to infer certain parameters in the model, another can be more appropriate to predict the future behavior of the system. These goals will eventually determine what outputs should be observed or computed and which criteria are most suited to judge a good experiment.

The Bayesian analysis aims to deduce from the prior pdf, through the Bayes theorem, the posterior pdf adjusting the parameter value by a probabilistic judgment in the light of the uncertainty of the data collected from a given sample. The prior pdf plays an important role in the Bayesian framework since it represents all the relevant information from the past data. It is also possible to incorporate any knowledge or even subjective opinion that one can have on the problem studied.

Shannon’s information gain is generally considered as an evaluation criterion of Bayesian designs. It is an adequate measure to evaluate the relevance of a proposed experiment under uncertainty. However, estimating the expected information gain is a challenging task due to its expensive computational cost.
1.2 Objectives

In this work, we use the MLMC method to compute the expected information gain for experimental design under a Bayesian setting for parameter inference goals. We compare the performance of this method to the Laplace approximation (LA) as well as the direct Double Loop Monte Carlo (DLMC). In fact, we implement the direct MLMC and the MLMC combined with the LA. Although the latter approach is faster due to the simplification made by the LA, the former method gives good results even in cases where the LA fails, such as when the posterior pdf is no longer concentrated. This case is particularly important for small number of experiments.

In this thesis, we test our work on two examples: the first example is the design of sensor deployment for a Darcy flow problem governed by a one-dimensional Poisson equation. We place the sensors in the locations where the pressure is measured and we model the conductivity field as a piecewise constant random vector with two parameters. The second one is chemical Enhanced Oil Recovery (EOR) core flooding experiment assuming homogeneous permeability. We measure the cumulative oil recovery, from a horizontal core flooded by water, surfactant and polymer, for different injection rates. The model parameters consist of the endpoint relative permeabilities, the residual saturations and the relative permeability exponents for the three phases: water, oil and microemulsions.

1.3 Related Work

Depending on the goal of the experiment, a wide range of criteria are proposed in the literature such as A-optimality and D-optimality criteria [1]. These criteria can be classified into Bayesian and non-Bayesian design criteria. For a thorough review of Bayesian experimental design criteria, see Chaloner and Verdinelli [2].
In the current work, we consider D-optimal Bayesian experimental design. This notion was first introduced by Lindley in [3] where the author considered the expected information gain as a utility function to evaluate the goodness of an experiment. One major obstacle in estimating this kind of utility function is that its evaluation may involve a high-dimensional integral. Conventional efficient approaches like sparse quadratures or double Monte Carlo method are computationally too expensive in this case.

Long et al. presented in [4] an alternative approach that consists in using a LA for the integration of the posterior probability density function (pdf). They were able to obtain a closed-form approximation of the inner integral reducing the estimation problem to a single-loop integration.

1.4 Contribution

In the present work, we implement the direct MLMC as well as the MLMC method combined with the LA for the computation of the expected information gain for the case of two numerical examples and we compare them to standard methods such as the DLMC and the MC method combined with the LA. The main novelty of this work is the cost comparison between the various methods for different tolerances. We also introduce an alternative approach to the estimation of the expected information gain based on the direct MLMC that appears to be an interesting approach in the cases where the LA fails.

1.5 Thesis structure

This thesis is outlined as follows: we start by presenting a brief survey on the key concepts of Bayesian Experimental Design in Chapter 2. Their understanding is crucial for introducing the goals of the thesis. Chapter 3 and 4 then provide the
tools necessary for the estimation of the expected information gain. We present the LA technique in Chapter 3 and Multilevel Monte Carlo method in Chapter 4. The numerical results are presented in Chapter 5. We show the results related to two examples: the first related to the design of sensor deployment for a Darcy flow problem governed by a one-dimensional Poisson equation. The sensors are placed to measure the pressure. The conductivity field is modeled as a piecewise constant random vector with two parameters. The second one concerns the study of chemical Enhanced Oil Recovery (EOR) core flooding experiment. Under the homogeneous permeability assumption, we measure the cumulative oil recovery, from a horizontal core flooded by water, surfactant and polymer, using different injection rates. The model parameters are the endpoint relative permeabilities, the residual saturations and the relative permeability exponents for the three phases: water, oil and microemulsions. Finally, the thesis ends with a summary of the work done and suggestions for future work.
Chapter 2

Bayesian Optimal Experimental Design

2.1 Background

Experimental data are crucial in order to perform inference or prediction. Also, conducting experiments is a vital part of the learning process since quite often experimentation is performed to investigate a particular hypothesis. For instance, an experiment can be carried out to test the fairness of a dice or to examine the effects of a given substance on the body. However, not all data are equally helpful in reducing the uncertainty about the parameter of interest. In fact, the results and conclusions that can be drawn from the analysis of an experiment depend to a large extent on the manner in which data were collected. Thus, the systematic design of experiments that yield the most informative data is needed.

Experimental design saw the light in the agricultural field in 1940 where experiments were conducted to determine which factors have an impact on the crop growth via the use of factorial experimental designs [5, 6]. Few years later, it started to gain popularity in the chemical industry [7, 8]. It was thanks to Tagushi [9, 10] that the discipline was well formulated allowing its spread in many other scientific and engineering fields. Nowadays, experimental design has a wide range of
applications in fields like pharmaceutics, physics and geophysics [11, 12, 13].

The goal of the scientists, when conducting an experiment, is to maximize the information content about the process. The aim of experimental design is to help scientists achieve the maximal information gain by designing optimal experiments.

Two classes of experimental design exist: Bayesian and non-Bayesian. The latter treats the data as being parametrized by the unknown variable that has a deterministic but unknown value. The former models the unknown as a random variable taking into account the operator’s belief using Bayes theorem.

Classical methods for experimental design, such as factorial or central composite designs, have received criticism over the recent years because usually the optimization criterion used, e.g. minimal variance of the parameters estimator, does not reflect quite often the objectives of the experiment [14]. Another issue to take into consideration is that, although most of these models are efficient in the case of linear or linearized models, they fail when facing complex nonlinear models. Unlike Bayesian experimental design, these techniques do not represent any unifying framework.

In many engineering fields, prior knowledge about the process or system of the study exists. Bayesian Experimental design can make use of this knowledge in order to reduce the number of needed experiments resulting in saving resources and time. In fact, decisions need to be taken before data collection since the experiment’s resources are often limited.

Kiefer and Wolfowitz [15, 16] were among the first authors to formulate a theory of optimal design. An optimal design problem arises whenever the scientist has the ability to select among many initial configurations for the experiment. In their works, they showed that, to overcome the dependence of exact optimal design on the sample size, the design needs to be formulated as a probability measure on the design space. They also introduced the notion of alphabetic optimality; a set of
criteria for experiment evaluation. The most widely accepted optimality criteria are \( c, A, D, E \) and \( G \) optimality.

The objective of an experiment is a crucial factor in determining how good an experiment is and which criterion is more suitable for the study. For instance, if we want to infer model parameters, then a good experiment minimizes the parameter variance. However, if we want to predict a quantity of interest accurately, then a good experiment minimizes the posterior predictive variance.

To explain briefly the difference between these criteria, we consider the example of the least squares estimate of the regression coefficients. Table 2.1 presents the objective function for the minimization problem associated with each criterion in this case. For further details about the optimality criteria, see [2].

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Objective function</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-optimality</td>
<td>Trace of the inverse of the information matrix</td>
</tr>
<tr>
<td>D-optimality</td>
<td>Determinant of the inverse of the information matrix</td>
</tr>
<tr>
<td>E-optimality</td>
<td>Maximum eigenvalue of the information matrix.</td>
</tr>
<tr>
<td>c-optimality</td>
<td>Variance of the least squares estimate of a linear combination of the regression coefficients</td>
</tr>
<tr>
<td>G-optimality</td>
<td>Variance of the average of a linear combination of the regression coefficients</td>
</tr>
</tbody>
</table>

D-optimal experimental design criterion is often considered in the literature when the goal of the experiment is to infer parameters whereas when the experiment aims to predict future values, G-optimality criterion is rather used.

Chaloner’s ideas [17] presented a significant milestone in the Bayesian optimal design theory foundation. Her contribution was important in the development of a robust theory for the case of linear design. However, the Bayesian optimality for non-linear models remains a challenging and active area of research. Bayesian experimental design framework is introduced in the next section.
2.2 Framework

Bayesian experimental design is a powerful method that offers significant advantages compared to classical designs. This approach incorporates prior knowledge (previous studies, literature, expert opinions) about the process into the design. Another advantage is that the number of experiments is arbitrary unlike the case of some classical design methods such as factorial design where the number of experiments is an integer power of two.

Designing an experiment is a very delicate framework that usually contains several steps. First step is to define clearly the objectives of the experiment. The next step is to determine the response and the factors involved in the experiment. Determining the number of possible repetitive experiments is the next stage in the process. This is very important since the scientist needs to take into account the time and resource constraints. After these preliminary but crucial steps comes the stage of the optimal design choice.

A unifying theory for Bayesian optimal experimental design was proposed by Lindley [18]. In the pursuit of determining the best possible design given the goals and constraints of the experiment, a utility function is defined. This kind of function measures how well an experiment did to accomplish the purpose behind it. The best design is thus the one that maximizes this function. A well-known example of a utility function quite often used in Bayesian experimental design is the Shannon information gain, see Stone, DeGroot and Bernardo [19, 20, 21]. An alternative approach to this is to define a loss function such as quadratic loss instead of a utility function. In this case, the aim is to minimize the loss function in order to find the optimal design.

Once the design is chosen, the experiment is performed and then comes the step of data analysis. This leads to the last part of this framework that is the summary of conclusions and recommendations.

It is worth mentioning that, in some cases, the experiment may have different
purposes and no single utility function can capture such goals. Defining a weighted utility function is an efficient way to deal with these situations. Also, a design considered optimal for inference purposes is not necessarily optimal for predicting future values of a given quantity of interest.

For a given experiment, we consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\Omega$ is the sample space, set of the potential values of the response $y$, $\mathcal{F}$ is the $\sigma$-field of measurable events and $\mathbb{P}$ is the probability measure.

Let $\theta: \Theta \mapsto \mathbb{R}^n$ be the uncertain parameters whose values to be inferred with a prior distribution $p(\theta)$, $y \in \mathbb{R}^m$ the set of data and $\xi \in \mathbb{R}^d$ the experimental set-up. The data model we are considering in this study is:

$$y_i = g(\theta_i, \xi) + \epsilon_i, \quad (2.1)$$

where $g(\theta_i, \xi)$ is the output of the computational model and the additive errors $\epsilon_i$ are independent and identically distributed (i.i.d) zero mean Gaussian errors with covariance matrix $\Sigma_\epsilon$. In general, the covariance matrix of the errors $\Sigma_\epsilon$ depends on the parameters $\theta$ as well as the experimental set-up $\xi$. However, we assume that such dependence does not exist in this work.

The mechanism of data collection entails various sources of uncertainties such as the experimental device, the experiment conditions, the computational model, etc. The addition of an error term to the computational model accounts for the uncertainty coming from these sources. In this case, we assume that the experimenter has control over the experimental set-up $\xi$ and thus the only uncertainty comes from the model parameters.

To illustrate these concepts, we consider an example from the astronomy field. A typical experiment consists of determining the times where to observe a given star in order to determine the orbit of a planet. In this case, $y$ represents the radial
velocity, $\theta$ is the time slots when the observations will take place which represents the experimental set-up and the parameters are the orbital parameters such as the period $\tau$, the eccentricity $e$ and the velocity amplitude $K$ of Keplerian velocity.

Bayesian experimental design is based on the well-known Bayes theorem:

$$p(\theta|y, \xi) = \frac{p(y|\theta, \xi)p(\theta|\xi)}{p(y|\xi)},$$  \hspace{1cm} (2.2)

where:

- $p(\theta|\xi)$ is the prior probability density function.
- $p(y|\theta, \xi)$ is the likelihood probability density function.
- $p(\theta|y, \xi)$ is the posterior probability density function.
- $p(y|\xi)$ is the evidence probability density function.

The prior distribution of $\theta$ can be interpreted as personal belief or prior knowledge of the scientist on the parameters before the data is simulated or collected. After gathering the data, this knowledge/belief is updated through the likelihood function. The posterior probability density function (pdf) can then be seen as the new belief since it incorporates both the experimental data as well as the prior distribution. Therefore, Bayes theorem can be regarded as an updating rule.

The denominator in Bayes theorem, the evidence pdf, is a normalizing factor that characterizes the prior data predictive density that defines all possible data that could be observed.

In most cases, the prior knowledge or belief that the experimenter has on the process or system needs to be used for constructing a prior density. This density must be either assumed or calculated. A well-known method to construct prior densities, when the experimenter disposes of incomplete information, is the maximum entropy principle. For instance, the uniform distribution is the maximum entropy distribution...
when the available information is the upper and lower bounds. If the experimenter has information on the mean and the variance, then the Gaussian prior would be the distribution obtained by means of the maximum entropy principle. For further details about this method, the reader is directed to [22].

2.3 Shannon information gain

In this study, the goodness indicator will be based on the Kullback-Leibler divergence ($D_{KL}$). The Kullback-Leibler divergence measures the relative entropy.

The entropy of a continuous random variable $\theta$ is defined as:

$$ h(\theta) = -\int_{\Theta} p(\theta) \log (p(\theta)) d\theta. \quad (2.3) $$

The entropy measures the degree of disorder or lack of information carried by the pdf of $\theta$. Low entropy indicates that few data are needed to estimate the true value of the unknown quantity whereas a high entropy means that, in order to determine the true value, a lot of data need to be collected.

The Kullback-Leibler divergence is quite used to compare the amount of information of two different pdfs, say $p_1(\theta)$ and $p_2(\theta)$, that represent the same unknown parameter $\theta$:

$$ D_{KL} = \int_{\Theta} p_1(\theta) \log \left( \frac{p_1(\theta)}{p_2(\theta)} \right) d\theta. \quad (2.4) $$

This quantity can be viewed as a penalty when using $p_2(\theta)$ to characterize the uncertain parameter when $p_1(\theta)$ is the right representation.

To evaluate the goodness of an experiment, it is natural to use a utility function that measures the amount of knowledge before and after performing the experiment. A typical example of a utility function is the Kullback-Leibler divergence between the
posterior and prior pdfs:

\[
D_{KL} = \int_{\Theta} p(\theta|y, \xi) \log \left( \frac{p(\theta|y, \xi)}{p(\theta)} \right) d\theta. \tag{2.5}
\]

A large \( D_{KL} \) can be interpreted as a sign that the data \( y \) have increased information of the belief of \( \theta \) by a large amount. Thus, these data are more informative with respect to the experimental goal.

Since the expression of the \( D_{KL} \) depends on the data \( y \), an unknown quantity at the design stage, then we can define the expected utility function by taking an expectation over \( Y \) to obtain:

\[
I(\xi) = \int_{\mathcal{Y}} \int_{\Theta} p(\theta|y, \xi) \log \left( \frac{p(\theta|y, \xi)}{p(\theta)} \right) d\theta p(y|\xi) dy. \tag{2.6}
\]

The last quantity is known as the expected information gain, a notion introduced by Shannon [23] and Lindley [3]. Experimental design is the only case where averaging over the sample space, an unknown quantity at the design stage, is meaningful.

An equivalent criterion to the expected information gain is the difference of entropy between the prior and the posterior densities defined as:

\[
- \Delta h = h(\theta) - h(\theta|y) = - \int_{\Theta} p(\theta) \log(p(\theta)) d\theta - \int_{\mathcal{Y}} p(y) h(\theta|Y = y, \xi) dy. \tag{2.7}
\]

The second term of equation (2.7) is constant for any experiment since it depends only on the prior distribution. Thus, the first term is relevant when comparing different designs. The proof of this equivalence is included in Appendix D.

### 2.4 Examples

In this section, three examples are briefly discussed.
2.4.1 Example 1: A linear design problem

An example of linear design problem can be to investigate the effect of the pressure $P$ and the temperature $T$ on the quality of the product $y$. The model in this case writes:

$$y = \theta_0 + \theta_1 P + \theta_2 T + \epsilon. \quad (2.8)$$

The experimental set-up in this case is $\xi = (P, T)^T$, $\theta = (\theta_0, \theta_1, \theta_2)^T$ is the regression parameters vector and $\epsilon$ is an additive zero mean Gaussian noise.

2.4.2 Example 2: A one-dimensional nonlinear design problem

The following model has been studied in [4]:

$$y(\theta, \xi) = \theta^3 \xi^2 + \theta \exp(-|0.2 - \xi|) + \epsilon, \quad (2.9)$$

where the noise has a Gaussian distribution with zero mean and variance equal to $10^{-2}$. The parameter $\theta$ has a uniform prior $U(0, 1)$.

Plot 2.1 shows the posterior pdf of the parameter for different values of the number of repetitive experiments $N_e$ given a value of $\xi = 0.2$. As $N_e$ increases, the posterior pdfs of $\theta$ concentrates around the value $\theta_i = 0.5$.

2.4.3 Example 3: Another nonlinear design problem

This example originally was introduced in a PhD thesis done by Button in Texas A&M (1979). The experiment was then studied by Atkinson et al. [24]. The aim of this experiment is to investigate the bioavailability, i.e. the quantity of a substance required to have an active effect when inserted into the body.
The experiment consisted in injecting the subject of the study, a number of horses, with 15mg/kg of theophylline as aminophylline by means of intragastric procedure. After injection, blood samples were taken at different times $t$ to measure the drug concentration $y$. The model used in this study is given by:

$$y = \theta_2(e^{-\theta_0 t} - e^{-\theta_1 t}) + \epsilon.$$  \hspace{1cm} (2.10)

Here, $\theta = (\theta_0, \theta_1, \theta_2)^T$ is the unknown vector parameter.

The design problem consisted in determining the times at which the blood samples are drawn. Atkinson et al. [24] studied the efficiency of the design proposed in Button’s thesis that was made up of 18 design points and one measurement was taken each time. Using different prior distributions, they looked for Bayesian optimal
experimental designs.

Analyzing the posterior pdf is crucial to evaluate the quality of an experiment. In general, the posterior pdf does not have an analytical expression. One way to tackle this problem is to evaluate the posterior at particular points that form the grid of the prior pdf support $\Theta$. This kind of technique becomes inefficient for high dimension of $\Theta$. In fact, the number of grid points grows in an exponential way with the dimension. In the next two chapters, we present two alternative methods: the Laplace approximation and the Multilevel Monte Carlo.
Chapter 3

Laplace approximation

3.1 Laplace approximation in Bayesian inference

The high dimensionality of the problem variables $\theta$ and $y$ makes it difficult to compute, in closed form, the expected information gain since often it involves a double-loop integration. The approach described in [4, 25] tries to approximate the posterior probability density function using a Gaussian approximation. They showed that this treatment accelerates significantly the estimation of the expected information gain.

The key idea behind this approach is to approximate the inner integral, in the case where the model parameters are determined by the experiment reducing the problem to a single-loop integration over the prior parameter space in order to estimate the expected information gain. The authors have also extended LA from the classical scenario, where a pdf has a simple dominant mode, to the situation where the pdf is characterized by a non-informative manifold [26].

The Laplace method [27] is an asymptotic method often used to approximate integrals. Some interesting improvements in the method’s implementation [28, 29], introduced in 1980, helped this technique gain popularity in the statistics community.

The use of the Laplace approximation in Bayesian inference comes as an attempt to reduce the computational time of methods such as Markov Chain Monte Carlo
(MCMC), maximum likelihood estimation or variational Bayes. The efficiency of LA in term of speed is highly dependent on the speed of the optimization problem used to find the mode of the posterior pdf. In general, the Laplace method is found to be faster than MCMC method since instead of characterizing the target pdf with enough samples, the method uses point estimates.

The idea of LA is to approximate the posterior pdf by a Gaussian one, simply by taking a second-order Taylor expansion around the mode of the posterior pdf logarithm. To guarantee a good quality of the approximation, the posterior pdf needs to be smooth and concentrated around its maximum to ensure the efficiency of the estimation. Such requirements are a must since the method considers that most of the contribution to the integral value comes from the maximum’s neighborhood. The Taylor expansion is very accurate when enough data points exist [28].

To investigate the quality of the approximation, we refer to the Bernstein-von Mises theorem [30, 31]. The theorem states that, if the data supplies enough amount of information, the posterior distribution is asymptotically normal, with mean and variance equal to the frequentist asymptotic statistics of the maximum likelihood estimator of the model used [32].

Let \( \{X_j\}_{j=1}^N \) are i.i.d random variables with a pdf \( f(x, \theta) \), \( \theta \in \Theta \subset \mathbb{R} \). We assume that \( \theta \) has a prior pdf and we denote by \( \hat{\theta}_N \) the maximum likelihood function of \( \theta \) based on \( \{X_j\}_{j=1}^N \). Then, we can state that:

\[
\sqrt{N}(\hat{\theta}_N - \theta_0) \text{ converges in distribution to } \mathcal{N}(0, \frac{1}{I(\theta_0)}) \text{ as } N \to +\infty
\]

where \( \hat{\theta}_N \to \theta_0 \) almost surely and \( I(\theta) = -\mathbb{E}_{\theta} \left[ \frac{\partial^2}{\partial \theta^2} \log f(x, \theta) \right] \).
3.2 Gaussian approximation of the posterior pdf

The use of the LA for estimating the expected information gain was introduced by Ryan [33] and also Long et al. [4]. Most of this chapter is based on the ideas of [4].

The posterior pdf of $\theta$ is given by Bayes theorem as:

$$p(\theta|y) \propto \prod_{i=1}^{N_e} \exp \left(-\frac{1}{2} r_i^T(\theta) \Sigma^{-1} r_i(\theta)\right) p(\theta),$$  \hspace{1cm} (3.1)

where $r_i(\theta) = g(\theta_i) + \epsilon_i - g(\theta)$, the residual of the $i^{th}$ measurement and $N_e$ is the number of repetitive experiments using the same experimental set-up.

In the rest of this chapter, for conciseness purposes, the experimental set-up $\xi$ is neglected since it appears in all the pdfs.

The Gaussian approximation of (3.1) leads to:

$$\tilde{p}(\theta|y) = (2\pi)^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} \exp \left(-\frac{1}{2} (\theta - \hat{\theta})^T \Sigma^{-1} (\theta - \hat{\theta})\right),$$  \hspace{1cm} (3.2)

where $\hat{\theta}$ is the maximum of the posterior pdf and $\Sigma$ is the inverse Hessian matrix of the negative logarithm of the posterior pdf evaluated at $\hat{\theta}$.

Long et al. [4] investigated how to determine the values of the parameters $\Sigma$ and $\hat{\theta}$. The negative logarithm of the posterior pdf is given by:

$$F(\theta) = -\log(p(\theta|y)) = \frac{1}{2} \sum_{i=1}^{N_e} r_i^T(\theta) \Sigma^{-1} r_i(\theta) - h(\theta) + C,$$  \hspace{1cm} (3.3)

where $h(\theta) = \log(p(\theta))$ and $C$ is a constant.

The second order Taylor expansion of $F(\theta)$ around $\hat{\theta}$ writes:

$$\tilde{F}(\theta) = F(\hat{\theta}) + \nabla F(\hat{\theta})(\theta - \hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^T \nabla^2 F(\hat{\theta})(\theta - \hat{\theta}) + O(||\theta - \hat{\theta}||^3).$$  \hspace{1cm} (3.4)
The Hessian of $F$, evaluated at $\hat{\theta}$, is given by:

$$\nabla^2 F(\hat{\theta}) = H(\hat{\theta})^T \Sigma^{-1}_\epsilon E_S + N_e J(\hat{\theta})^T \Sigma^{-1}_\epsilon J(\hat{\theta}) - \nabla^2 h(\hat{\theta}), \quad (3.5)$$

where $J$ and $H$ are the Jacobian and Hessian of $-g$ with respect to $\theta$ and $E_S$ is the sum of residuals $\{r_i\}_{i=1}^{N_e}$, i.e. $E_S = \sum_{i=1}^{N_e} r_i$.

Going back to (3.4), we should note that since $\hat{\theta}$ is the maximum of the posterior pdf then $\nabla F(\hat{\theta}) = 0$ and that $F(\hat{\theta})$ is simply a shifting constant that does not change the shape of the density. Thus, we can see that using this framework leads to a posterior pdf with a Gaussian shape as suggested in (3.2) where $\Sigma = (\nabla^2 F(\hat{\theta}))^{-1}$.

Long et al. [4] showed that the first term in (3.5) is bounded in probability, $H(\hat{\theta})^T \Sigma^{-1}_\epsilon E_S = O_P(\sqrt{N_e})$, i.e $\forall \epsilon > 0, \exists K = K(\epsilon)$ and $m_0 = m_0(\epsilon)$ such that:

$$P(|H(\hat{\theta})^T \Sigma^{-1}_\epsilon E_S| \leq K \sqrt{N_e}) \geq 1 - \epsilon, \forall N_e > m_0,$$ see Appendix C. Therefore, an approximate expression of $\Sigma$ is given by:

$$\Sigma \approx (N_e J(\hat{\theta})^T \Sigma^{-1}_\epsilon J(\hat{\theta}) - \nabla^2 h(\hat{\theta}))^{-1}. \quad (3.6)$$

In this study, we assume that the Hessian of $F$, evaluated at $\hat{\theta}$ is full rank without help from the prior, in other words, the model parameters are asymptotically determined by the experiment.

To find the mode of the posterior pdf, we need to minimize $F(\theta)$, i.e.

$$\hat{\theta} = \arg\min_{\theta} [F(\theta)]:$$

$$\hat{\theta} = \arg\min_{\theta} \left[ \frac{1}{2} N_e (g(\theta_t) - g(\theta))^T \Sigma^{-1}_\epsilon (g(\theta_t) - g(\theta)) + E_S^T \Sigma^{-1}_\epsilon (g(\theta_t) - g(\theta)) - h(\theta) \right].$$

Long et al. [4] derived a first order approximation of $\hat{\theta}$:

$$\hat{\theta} = \theta_t - (N_e J^T \Sigma^{-1}_\epsilon J + H^T \Sigma^{-1}_\epsilon E_S - \nabla^2 h(\theta_t))^{-1} J^T \Sigma^{-1}_\epsilon E_S + O_P\left(\frac{1}{N_e}\right). \quad (3.7)$$
When $N_e \to +\infty$, we can state that, in this case, we have:

$$\hat{\theta} = \theta_t + O_P\left(\frac{1}{\sqrt{N_e}}\right).$$  \hfill (3.8)

The proof of this statement can be found in Appendix C

### 3.3 Laplace approximation for the expected information gain estimation

After explaining briefly how to compute the values of $\Sigma$ and $\hat{\theta}$, we turn our attention to determining a closed-form approximation for the Kullback Liebler divergence which we recall its expression:

$$D_{KL} = \int_\Theta \log \left( \frac{p(\theta|y)}{p(\theta)} \right) p(\theta|y)d\theta. \hfill (3.9)$$

Using (3.2), we can write the ratio of the logarithm of the posterior and the prior as:

$$\log \left( \frac{p(\theta|y)}{p(\theta)} \right) = \log \left( \frac{p(\theta|y)}{\tilde{p}(\theta|y)} \times \frac{\tilde{p}(\theta|y)}{p(\theta)} \right)$$

$$= \log \left( \frac{p(\theta|y)}{\tilde{p}(\theta|y)} \right) + \log \left( \frac{\tilde{p}(\theta|y)}{p(\theta)} \right)$$

$$= \log \left( \frac{p(\theta|y)}{\tilde{p}(\theta|y)} \right) + \log(\tilde{p}(\theta|y)) - \log(p(\theta)),$$

and thus, we get:

$$\log \left( \frac{p(\theta|y)}{p(\theta)} \right) = \epsilon_{LA} - \frac{1}{2} \log((2\pi)^d|\Sigma|) - \frac{1}{2}(\theta - \hat{\theta})^T \Sigma^{-1}(\theta - \hat{\theta}) - h(\theta). \hfill (3.10)$$

where $\epsilon_{LA} = \log \left( \frac{p(\theta|y)}{\tilde{p}(\theta|y)} \right)$ is the logarithm error resulting from approximating the posterior pdf by a Gaussian one.
Using (3.9) and (3.10), we can write:

\[
D_{KL} = \int_{\Theta} \left[ -\frac{1}{2} \log((2\pi)^d |\Sigma|) - \frac{1}{2} (\theta - \hat{\theta})^T \Sigma^{-1} (\theta - \hat{\theta}) - h(\theta) + \epsilon_{LA} \right] \tilde{p}(\theta | y) d\theta + \epsilon_{int},
\]

where \( \epsilon_{int} = \int_{\Theta} \log(\tilde{p}(\theta | y)) (p(\theta | y) - \tilde{p}(\theta | y)) d\theta. \)

In [4], the authors showed the following estimates:

- \( \epsilon_{int} = O_P(\frac{1}{N^e}) \),
- \( \int_{\Theta} \epsilon_{LA} \tilde{p}(\theta | y) d\theta = O_P(\frac{1}{N^e}) \),
- \( \int_{\Theta} h(\theta) \tilde{p}(\theta | y) d\theta = h(\hat{\theta}) + \frac{1}{2} tr(\Sigma \nabla \nabla h(\hat{\theta})) + O_P(\frac{1}{N^e}). \)

This leads to the following closed-form approximation for the Kullback Liebler divergence:

\[
D_{KL} = -\frac{1}{2} \log((2\pi)^d |\Sigma|) - \frac{d}{2} - h(\hat{\theta}) - \frac{1}{2} tr(\Sigma \nabla \nabla h(\hat{\theta})) + O_P(\frac{1}{N^e}). \tag{3.11}
\]

Going back to our original goal of estimating the expected information gain, we can write:

\[
I = \int_{\gamma} D_{KLP}(y) dy
\]

\[
= \int_{\gamma} \int_{\Theta} D_{KLP}(y|\theta_t) p(\theta_t) dy d\theta_t
\]

\[
= \int_{\gamma} \int_{\Theta} \left[ \frac{1}{2} \log((2\pi)^d |\Sigma|) - \frac{d}{2} - h(\hat{\theta}) - \frac{1}{2} tr(\Sigma \nabla \nabla h(\hat{\theta})) \right] p(y|\theta_t) p(\theta_t) dy d\theta_t + O(\frac{1}{N^e}).
\]

Using (3.7), we can write:

\[
I = \int_{\Theta} \left[ \frac{1}{2} \log((2\pi)^d |\Sigma|) - \frac{d}{2} - h(\theta_t) \right] p(\theta_t) d\theta_t + O(\frac{1}{N^e}). \tag{3.12}
\]

Form (3.6), we have \(|\Sigma| = O(N^{-d}_e)\) where \(d\) is the dimension of \(\theta\). Therefore, using the expression (3.12), can see that the expected information gain grows as
\log(N_e).

In order to compute the expected information gain, using the expression (3.12), the LA can be combined with other numerical techniques such as the Monte Carlo methods or the sparse grid quadrature. Next, in section 4.2.3 of chapter 4, we describe how to use the MLMC to compute the integral (3.12).
Chapter 4

Multilevel Monte Carlo

4.1 Standard Monte Carlo

4.1.1 Monte Carlo for numerical integration

Monte Carlo (MC) method is a widely used numerical tool for studying the characteristics of systems that incorporate random parameters. The key idea is to draw samples of the random variable (RV) and to use a mathematical model, that describes the dependency of the quantity of interest on the random component, in order to compute the statistics, e.g. the mean, the variance, the confidence interval, of the quantity of interest.

In many engineering applications, the aim is to compute the expected value of a random variable $Y$, $\mathbb{E}[Y]$. Let $\{Y(\omega)\}_{j=1}^{N}$ be i.i.d. samples drawn from the distribution $\rho_Y$ of $Y$ where $N$ is chosen sufficiently large to control the statistical error, $\varepsilon_N = \mathbb{E}[Y] - I_N(Y)$. Here, $I_N(Y)$ is the standard MC estimator given by:

$$ I_N(Y) = \sum_{j=1}^{N} Y(\omega_j). \tag{4.1} $$

The rate of convergence of the MC method is given by the Central Limit theorem (CLT):
Theorem 1. Let \( \{\xi_j\}_{j=1}^N \) be a sequence of i.i.d random variables such as \( \mathbb{E}[\xi_j] = 0 \) and \( \mathbb{E}[\xi_j^2] = 1 \). Then:
\[
\sum_{j=1}^N \frac{\xi_j}{\sqrt{N}} \rightarrow \nu,
\]
where \( \nu = N(0,1) \) and \( \rightarrow \) denotes the weak convergence, i.e. for all bounded and continuous functions \( g \), we have:
\[
\mathbb{E} \left[ g\left( \sum_{j=1}^N \frac{\xi_j}{\sqrt{N}} \right) \right] \rightarrow \mathbb{E}[g(\nu)].
\]

Given a constant \( 0 < \alpha \ll 1 \), the CLT gives us an idea on how the statistical error behaves for large \( M \):
\[
P\left( |\varepsilon_N| \leq \frac{\sigma C_\alpha}{\sqrt{N}} \right) \approx 1 - \alpha. \tag{4.2}
\]
where \( \sigma = 1 \) in this case and \( C_\alpha \) is a constant that depends on the value of \( \alpha \).

Thus, the statistical error of the Monte Carlo estimator is \( \mathcal{O}_P(N^{-\frac{1}{2}}) \).

We illustrate this result in figure 4.1. We plot the computational error when using MC to estimate the integral:
\[
I = \frac{1}{(e-1)^{20}} \int_{[0,1]^{20}} \exp(\sum_{i=1}^{20} x_i)dx_1dx_2\ldots dx_{20}.
\]

We note that the slope of the curve has a rate of the order of \( N^{-\frac{1}{2}} \) which coincides with the theoretical behavior.
4.1.2 Monte Carlo for the expected information gain estimation

4.1.2.1 Monte Carlo coupled with Laplace approximation

In chapter 3, LA was used to reduce the expected information gain expression to a single integration over the parameter space. To evaluate this integral, different numerical techniques, such as sparse quadrature or MC, can be used.

Using the discussion in the previous section, we explain briefly how to use standard MC to estimate the expected information gain using Algorithm 1.

**Algorithm 1** Algorithm to compute the expected information gain using (MC+LA)

1: Sample $N$ times the random vector $\theta_j \in \mathbb{R}^d$, $\forall j = 1, \ldots, N$
2: Compute the values $f(\theta_j) = -\frac{1}{2} \log((2\pi)^d|\Sigma|) - \frac{d}{2} - h(\theta_j)$, $\forall j = 1, \ldots, N$
3: Compute the MC estimate $I_N = \frac{1}{N} \sum_{j=1}^{N} f(\theta_j)$
4: Estimate the statistical error $\varepsilon_N = \mathbb{E}[D_{KL}] - I_N(D_{KL})$ using for instance the CLT 4.2.
5: If the statistical error is not small enough, then increase $N$ and go to step 1
For a given tolerance $TOL$, the error of the MC+LA is estimated by:

$$\varepsilon_{TOL} \approx C_1 h_L^{\eta} + \frac{C_2}{\sqrt{N}} + \frac{C_3}{N_e}$$  \hspace{1cm} (4.3)$$

where $\{C_i\}_{i=1}^3$ are constants to be estimated and $\eta$ an exponent that depends on the numerical method used for the discretization.

Assuming that the cost per sample is $O(h_L^{-\gamma})$, then the corresponding work, in the case when $TOL \geq \frac{C_3}{N_e}$, is given by:

$$Work = O(TOL^{-(2+\frac{\gamma}{\eta})})$$ \hspace{1cm} (4.4)$$

Although the rate of Monte Carlo is independent of the dimension of the problem, $d$, it is relatively slow. In order to improve this convergence, quite often variance reductions techniques, such as antithetic variates or control variates, are used. The key idea behind these methods is to find a random variable $Z$ having the same mean as $Y$ but a smaller variance. For more details about these techniques, see [34].

### 4.1.3 Double Loop Monte Carlo

In situations where the LA assumptions, such as the concentration of measures, are not satisfied, an alternative approach is to compute the expected information gain using Double Loop Monte Carlo (DLMC).

The MC estimator for the expected information gain, (2.6), is given by:

$$I \approx \frac{1}{N} \sum_{n=1}^{N} \log \left( \frac{p(y_n|\theta_n)}{p(y_n)} \right).$$ \hspace{1cm} (4.5)$$

Here, we sample the parameter $\theta_n$ from the prior $p(\theta)$ and the data $y_n$ are drawn from the likelihood $p(y|\theta_n)$. 
To evaluate the evidence, \( p(y_n) \), we use the MC method once more:

\[
p(y_n) \approx \frac{1}{M} \sum_{m=1}^{M} p(y_n|\theta_{n,m}). \tag{4.6}
\]

Thus, the DLMC estimator is given by:

\[
I \approx \frac{1}{N} \sum_{n=1}^{N} \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M} \sum_{m=1}^{M} p(y_n|\theta_{n,m})} \right). \tag{4.7}
\]

where \( N \) and \( M \) are the number of samples for the outer and inner loop respectively. In [33], Ryan investigated the properties of the estimator (4.7). In fact, he showed, with a detailed proof, that:

- The variance of such estimator is \( O(N^{-1}) \).
- The bias is \( O(M^{-1}) \).

Algorithm 2 explains the general procedure to implement the DLMC method to compute the expected information gain.

**Algorithm 2** Algorithm to compute the expected information gain using DLMC

1: Generate a large sample \( \{\theta_1, \ldots, \theta_{N_0}\} \) of size \( N_0 \) from \( p(\theta) \)
2: Draw an index set of size \( N \leq N_0 \) random sample of the integers 1 to \( N_0 \) to get samples for the outer loop \( \{\theta_1, \ldots, \theta_N\} \)
3: Draw the data sample, \( \{y_n\}_{n=1}^{N} \), from the likelihood \( p(y|\theta_n) \)
4: Generate index sets as \( N \) independent size \( M \leq N_0 \) random samples of the integers 1 to \( N_0 \) to get samples for the inner loop \( \{\theta_{1,i}, \ldots, \theta_{N,i}\} \) for \( i = 1, \ldots, M \)
5: Compute the estimator using (4.7)

In general, the cost of the DLMC method is \( O(N_0 + N \times M) \). The linear cost comes from the sampling of \( \theta \) and thus this cost matters only if the model is quite simple. In our case, this linear cost is negligible and therefore the cost is \( O(N \times M) \).

The DLMC method can become quickly very expensive, since for a given tolerance \( TOL \), the work of the method is \( O(TOL^{-3}) \). Reusing samples between the inner and outer MC loops can be efficient in order to reduce the required number of
forward model evaluations and thus improve the computational cost. According to [35], this procedure reduces the cost for a given design from $O(N \times M)$ to $O(N)$. Although this treatment increases the bias of the estimator, this effect was shown to be very small in the numerical example studied in [36]. This cost analysis is done without taking into account the bias coming from the discretization. In our work, we need to consider such component. Thus, the approximate expression of the error of the DLMC is:

$$
\varepsilon_{TOL} \approx C_1 h_{TOL}^n + \frac{C_4}{\sqrt{N}} + \frac{C_5}{M}
$$

(4.8)

where $C_4$ and $C_5$ are constants to be estimated.

Finally, the work of this method is:

$$
Work = O(TOL^{-(3+\frac{2}{5})})
$$

(4.9)

### 4.2 Multilevel Monte Carlo

#### 4.2.1 Multilevel Monte Carlo for numerical integration

In many applications, for instance in the case of problems involving stochastic partial differential equations (SPDEs) or random input partial differential equations (RPDEs), the cost of generating samples is expensive and thus MC method becomes very slow. In recent years, a new method, Multilevel Monte Carlo (MLMC) has been developed to overcome this issue. In fact, most of the simulations are done with low accuracy at low cost whereas only a few simulations are performed with high accuracy at high cost.

The idea of MLMC was first introduced in the context of parametric integration when Heinrich et al. [37, 38] used this technique to estimate $\mathbb{E}[f(x, \lambda)]$ where $x$ is
finite-dimensional random vector and $\lambda$ is a parameter. Similar ideas were used by Kebaier [39] to approximate weak solutions of SDEs in finance using a two-level MC estimator.

It was thanks to Giles [40] who has generalized these ideas to what is now known as MLMC using a similar framework to control variates. In fact, the method uses different approximations of the same RV and thus the sampling comes from several estimates of the quantity of interest.

Let $u$ denote the solution of a SPDE or a RPDE. We seek to compute the expectation of a real valued function, $\mathbb{E}[P(u)]$, while controlling the statistical error. Let $P_l$ be the corresponding numerical approximation on level $l$ using a discretization $\{h_l\}_{l=0}^L$.

The expected value of the approximation on the finest level is given by:

$$
\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{l=1}^L \mathbb{E}[P_l - P_{l-1}].
$$

Equation (4.10) expresses the expectation of the finest approximation as the sum of the expectation on the coarsest level plus a telescopic sum that represents the expectation of the difference between approximations using different step sizes.

The idea behind MLMC method is to approximate, in an independent way, these expected values using standard MC. On the coarse grid, the variance of the respective difference is big so the number of samples needed per level but the cost of computations is less expensive. However, on the fine grid, the value of the variance of the respective difference becomes smaller so the optimal number of samples per level but the cost becomes greater.

The MLMC estimator is given by:

$$
Y = \sum_{l=0}^L Y_l = \sum_{l=0}^L \frac{1}{N_l} \sum_{n=1}^{N_l} (P_l(\omega_{l,n}) - P_{l-1}(\omega_{l,n})).
$$

(4.11)
where $P_{-1} = 0$ and the quantity $(P_l(\omega_{l,n}) - P_{l-1}(\omega_{l,n}))$ results from the same random sample but computed on different levels $l$ and $l - 1$.

Let $V_l = Var[P_l - P_{l-1}]$, then the statistics of the MLMC estimator are:

- $Var[Y] = \sum_{l=0}^{L} V_l N_l^{-1}$.

Let $C_0, C_l$ denote respectively the cost of one sample of $P_0$ and $(P_l - P_{l-1})$ and $V_0, V_l$ to be the variance of $P_0$ and $(P_l - P_{l-1})$, then the total cost of MLMC method is $\sum_{l=0}^{L} N_l C_l$.

In order to determine the optimal number of samples per level, $N_l$, we introduce the optimization problem that consists in minimizing the total cost for a fixed variance:

$$\text{Min} \quad \sum_{l=0}^{L} N_l C_l \quad \text{Subject to} \quad \sum_{l=0}^{L} N_l^{-1} V_l = \frac{TOL^2}{2}$$

We introduce the Lagrange multiplier, $\mu^2$, to reformulate the problem into minimizing the quantity:

$$\sum_{l=0}^{L} N_l C_l + \mu^2 \left( N_l^{-1} V_l - \frac{TOL^2}{2} \right).$$

Solving the optimization problem yields $N_l = \mu \sqrt{V_l C_l^{-1}}$. To achieve a total variance of $TOL^2$, it is required that $\mu = TOL^{-2} \sum_{l=0}^{L} \sqrt{V_l C_l}$. Therefore, the overall computational cost is:

$$C = TOL^{-2} \left( \sum_{l=0}^{L} \sqrt{V_l C_l} \right)^2. \quad (4.12)$$

It is crucial to investigate how the product $V_l C_l$ behaves with the level $l$, i.e. to compare the increase of the cost to the decrease of the variance per level. We have three situations:
- $V_l C_l$ increases with the level: the dominant contribution to the cost comes from the finest level and the cost is $C \approx TOL^{-2} V_L C_L$,

- $V_l C_l$ decreases with the level: most of the contribution comes from the coarsest level and the cost is $C \approx TOL^{-2} V_0 C_0$,

- $V_l C_l$ does not vary with the level: the total cost is then $C \approx TOL^{-2} L^2 V_0 C_0 = TOL^{-2} L^2 V_L C_L$.

If we assume that the cost of computing $P_l$ is similar to the one of computing $(P_l - P_{l-1})$ and that $Var[P_L] \approx Var[P_0]$, then the cost of standard MC is approximately $TOL^{-2} V_0 C_0$. Therefore, the cost reduction factor compared with MLMC is given by:

- $\frac{V_L}{V_0}$, if $V_l C_l$ increases with the level $l$,

- $\frac{C_0}{C_L}$, if $V_l C_l$ decreases with the level $l$,

- $L^2 \frac{V_l}{V_0} = L^2 \frac{C_0}{C_L}$, if $V_l C_l$ is constant with respect to the level $l$.

In [41], Giles states a fundamental theorem regarding Multilevel Monte Carlo complexity. In this theorem, he mentions the conditions needed to control the mean squared error. The statement of the theorem is the following:

**Theorem 2.** Let $P$ denotes a RV and let $P_l$ denotes the corresponding level $l$ numerical approximation. If there exist independent estimators $Y_l$ based on $N_l$ MC samples, each with expected cost $C_l$ and variance $V_l$, and positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that $\alpha \geq \frac{1}{2} \min(\beta, \gamma)$ and

i. $|\mathbb{E}[P_l - P]| \leq c_1 2^{-\alpha l}$

ii. $\mathbb{E}[Y_l] = \begin{cases} 
\mathbb{E}[P_0] & \text{if } l = 0 \\
\mathbb{E}[P_l - P_{l-1}] & \text{if } l > 0 
\end{cases}$

iii. $V_l \leq c_2 2^{-\beta l}$
iv. \( C_i \leq c_3 2^n \)

then there exists a positive constant \( c_4 \) such that for any \( TOL < e^{-1} \) there are values \( L \) and \( N_i \) for which the multilevel estimator \( Y = \sum_{l=0}^{L} Y_l \) has a mean-square-error (MSE) with bound:

\[
MSE := \mathbb{E}[(Y - \mathbb{E}[P])^2] < TOL^2,
\]  

with a computational complexity \( C \) with bound:

\[
\mathbb{E}[C] \leq \begin{cases} 
    c_4 TOL^{-2} & \text{if } \beta > \gamma \\
    c_4 TOL^{-2} (\log TOL)^2 & \text{if } \beta = \gamma \\
    c_4 TOL^{-2 - (\gamma - \beta)/\alpha} & \text{if } \beta < \gamma 
\end{cases}
\]

Theorem 2 generalizes the original theorem in [40]. The proof of the theorem exists in [42] with the small change of considering expected costs to take into consideration the randomness of the cost of sampling in some cases. The MSE expression can be seen as the sum of two terms, the variance of the estimator and the squared bias:

\[
MSE = Var[Y] + (\mathbb{E}[Y - P])^2.
\]  

To ensure that the MSE is less than \( TOL^2 \), it is sufficient to ensure that both \( Var[Y] \) and \((\mathbb{E}[Y - P])^2\) are both less than \( \frac{TOL^2}{2} \). The number of levels, \( L \), is chosen so that \((\mathbb{E}[Y - P])^2 < \frac{TOL^2}{2}\) whereas \( N_i \) is determined to ensure that \( Var[Y] < \frac{TOL^2}{2} \).

The computational cost expression is based on a comparison between \( \beta \) and \( \gamma \), three possible situations exist:

**First case:** \( \beta > \gamma \)

The dominant computational cost is on the coarsest levels where \( C_i = \mathcal{O}(1) \) and \( \mathcal{O}(TOL^{-2}) \) samples are needed to achieve the required accuracy.
Second case: \( \beta < \gamma \)

The dominant computational cost is on the finest levels because of condition (i. \( 2^{-\alpha L} = \mathcal{O}(TOL) \)) and hence \( C_L = \mathcal{O}(2^{\gamma L}) = \mathcal{O}(TOL^{-\gamma/\alpha}) \). If \( \beta = 2\alpha \), which is usually the best that can be achieved since typically \( Var[P_l - P_{l-1}] \) is similar in magnitude to \( \mathbb{E}[(P_l - P_{l-1})^2] \) which is greater than \( (\mathbb{E}[P_l - P_{l-1}])^2 \). In this case, since \( \beta = 2\alpha \), the total cost is then \( \mathcal{O}(TOL^{-\gamma/\alpha}) = \mathcal{O}(C_L) \), corresponding to \( \mathcal{O}(1) \) samples on the finest level which is the best that can be achieved.

Third case: \( \beta = \gamma \)

The computational effort and the contributions to the overall variance are spread approximately evenly across all the levels.

The MLMC algorithm is given below:

\begin{verbatim}
Algorithm 3 MLMC algorithm
1: Start with \( L = 2 \), and initial target of \( N_0 \) samples on levels \( l = 0, 1, 2 \)
2: while Extra samples need to be evaluated do
3: Evaluate extra samples on each level
4: Compute/update estimates for \( V_l, l = 0, \ldots, L \)
5: Define optimal \( N_l, l = 0, \ldots, L: N_l = [2TOL^{-2}\sqrt{V_l/C_l}\sum_{i=0}^{L} V_i/C_i] \)
6: Test for weak convergence: \( \frac{\mathbb{E}[P_l - P_{l-1}]}{2^{\alpha l}} < \frac{TOL}{\sqrt{2}} \)
7: If not converged, set \( L := L + 1 \) and initialize target \( N_L \)
8: end while
\end{verbatim}

The optimal expression of \( N_l \) is derived to ensure that \( Var[Y] < \frac{TOL^2}{2} \) and the weak convergence test is to check that \( (\mathbb{E}[Y - P])^2 < \frac{TOL^2}{2} \). In this case, the MSE is less than \( TOL^2 \). In our work, we use Giles implementation for the MLMC [41]. Recent improvements of MLMC ideas can be found in [43, 44]. In [43], an efficient and robust implementation of the MLMC is presented. The needed parameters are estimated on the fly while controlling the error in the weak sense. In [44], the authors generalize Theorem 2 and also derive precise constants for the asymptotic convergence of the MLMC in terms of input parameters.
4.2.2 Multilevel Monte Carlo for the expected information gain estimation

We now apply the idea of MLMC to compute the expected information gain. We start by recalling that the expression of the expected information gain using the DLMC is given by:

\[ I \approx \frac{1}{N} \sum_{n=1}^{N} \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M} \sum_{m=1}^{M} p(y_n|\theta_{n,m})} \right). \]  

(4.15)

Note that as explained in section 4.1.3, both \( M \) and \( N \) need to be sufficiently large in order to have a good accuracy, i.e. to control both the bias and the statistical error. However, a straightforward implementation will lead to a significant increase in the computational cost.

We can apply MLMC ideas to construct an efficient estimator based on the expression of (4.15). We define an increasing geometric sequence, \( M_l = M_0 \times 2^l \), for the choice of the number of inner samples and we determine the optimal number of outer samples using the expression detailed in Algorithm 3. In order to construct a low variance estimator, we can use a variance reduction technique, the antithetic variates [41]. The MLMC estimator is this case is defined as:

\[
Y_l = \frac{1}{N_l} \sum_{n=1}^{N_l} \left[ \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M_l} \sum_{m=1}^{M_l} p(y_n|\theta_{n,m})} \right) - \frac{1}{2} \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M_{l-1}} \sum_{m=M_{l-1}+1}^{M_l} p(y_n|\theta_{n,m})} \right) \right].
\]

(4.16)

We define:

\[ \psi(y_n, \theta_m) = p(y_n|\theta_m) \] and \( f \left( \frac{1}{M_l} \sum_{m=1}^{M_l} \psi(y_n, \theta_m) \right) = \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M_l} \sum_{m=1}^{M_l} p(y_n|\theta_{n,m})} \right). \]
Then, we can write:

\[
Y_l = \frac{1}{N_l} \sum_{n=1}^{N_l} \left[ f \left( \frac{1}{M_l} \sum_{m=1}^{M_l} \psi(y_n, \theta_m) \right) - \frac{1}{2} f \left( \frac{1}{M_{l-1}} \sum_{m=1}^{M_{l-1}} \psi(y_n, \theta_m) \right) \right] - \frac{1}{2} f \left( \frac{1}{M_{l-1}} \sum_{m=M_{l-1}+1}^{M_l} \psi(y_n, \theta_m) \right).
\]  

(4.17)

If \( f \) is twice differentiable, then using a second order Taylor expansion, see Appendix A, we find that:

\[
Y_l \approx -\frac{1}{8N_l} \sum_{n=1}^{N_l} f''(\mathbb{E}[\psi(y_n, \theta_m)]) (\Delta \psi_1^{(n)} - \Delta \psi_2^{(n)})^2.
\]

where \( \Delta \psi_1^{(n)} \) and \( \Delta \psi_2^{(n)} \) are the remainder terms of the Taylor expansion of the second and third term in (4.17):

\[
\frac{1}{M_{l-1}} \sum_{m=1}^{M_{l-1}} \psi(y_n, \theta_m) = \mathbb{E}[\psi(y_n, \theta_m)] + \Delta \psi_1^{(n)}.
\]

\[
\frac{1}{M_{l-1}} \sum_{m=M_{l-1}+1}^{M_l} \psi(y_n, \theta_m) = \mathbb{E}[\psi(y_n, \theta_m)] + \Delta \psi_2^{(n)}.
\]

Using the Chebyshev inequality, we have: \( \Delta \psi_i^{(n)} = \mathcal{O}_P(M_i^{-\frac{1}{2}}) \) for \( i = 1, 2 \) and thus, we get:

\[
\mathbb{E}[Y_l] = \mathcal{O}(M_i^{-1}).
\]

\[
V_i = \mathcal{O}(M_i^{-2}).
\]

In this work, we use a geometric sequence \( M_i = M_0 \times 2^i \) to define the number of inner samples. Therefore, we can write:

\[
\mathbb{E}[Y_l] = \mathcal{O}(2^{-i}).
\]

\[
V_i = \mathcal{O}(2^{-2i}).
\]
Using Giles MLMC theorem [40], this correspond to the first case ($\beta = 2, \gamma = 1$), therefore the MLMC estimation based on $Y_l$ defined in (4.16) has a complexity of $O(TOL^{-2})$.

The previous discussion assumes that the value of the expected information gain is computed using a fixed discretization. When taking into consideration the error contribution from the discretization, we can write the total error as:

$$\varepsilon_{TOL} \approx C_1 h_{TOL}^\eta + \sqrt{\sum_{l=0}^{L} \frac{V_l}{N_l}} + \frac{C_5}{M_L}$$

(4.18)

In this case, the optimal number of samples $N_l$ is computed to satisfy that $Var[Y] < \frac{TOL^2}{3}$ and the weak convergence test consists on checking that $(E[Y - P])^2 < \frac{TOL^2}{3}$. Therefore, the work of the method is $O(TOL^{-(2+\frac{\gamma}{2})})$.

4.2.3 LA combined with MLMC

Since the expected information gain can be written as $I = \mathbb{E}[D_{KL}]$, then we can use MLMC technique explained in Chapter 4 to compute this quantity. In order to accelerate this computation, we can use the expression (3.12).

In MLMC setting, we can write:

$$\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{l=1}^{L} \mathbb{E}[P_l - P_{l-1}],$$

(4.19)

where

$$P_l(\theta) = \frac{1}{2} \log((2\pi)^d |\Sigma_l(\hat{\theta})|) - \frac{d}{2} - h(\theta),$$

$$\Sigma_l(\hat{\theta}) \approx \left( N_\epsilon J_l(\hat{\theta})^T \Sigma^{-1}_\epsilon J_l(\hat{\theta}) - \nabla \nabla h(\hat{\theta}) \right)^{-1}.$$  

(4.20)

The expected information gain MLMC estimator can be written as follows:

$$I = \sum_{l=0}^{L} I_l = \sum_{l=0}^{L} \frac{1}{N_l} \sum_{n=1}^{N_l} (P_l(\theta_{l,n}) - P_{l-1}(\theta_{l,n})).$$

(4.21)
where $P_{-1} = 0$ and the quantity $(P_l(\theta_{t,n}) - P_{l-1}(\theta_{t,n}))$ results from the same random sample but computed on different levels $l$ and $l - 1$. Here, the level $l$ is associated to the level of mesh refinement in the discretization of the differential equation using for example the finite element method.

Now, we will try to analyze the cost and accuracy of this method. We are concerned with the behavior of $P_l$ with respect to the discretization, $h_l$, relative to the level $l$. Thus, we can focus the derivation on the quantity $\hat{P}_l$ defined as:

$$\hat{P}_l(\hat{\theta}) = \log(|\Sigma_l(\hat{\theta})|).$$  \hspace{1cm} (4.22)

Note that $\hat{P}_l$ contains the terms of $P_l$ that depends on the discretization. The other terms will not affect the general behavior with respect to $h_l$.

We assume that:

$$|\Sigma_l(\hat{\theta})| = T(\hat{\theta}) + C(\hat{\theta})h_l^n + o(h_l^n).$$  \hspace{1cm} (4.23)

where $\eta$ is a constant that depends on the forward problem. The relation 4.23 simply expresses the speed of convergence of $|\Sigma_l(\hat{\theta})|$ with respect to $h_l$. Here, $T$ is the true value of $|\Sigma_l(\hat{\theta})|$ and $C$ is a constant such that $\frac{C}{T}h_l^n \ll 1$.

$$\hat{P}_l(\hat{\theta}) = \log(|\Sigma_l(\hat{\theta})|)$$

$$= \log(T + Ch_l^n)$$

$$= \log(T) + \log(1 + \frac{C}{T}h_l^n)$$

$$\approx \log(T) + \frac{C}{T}h_l^n$$  \hspace{1cm} (4.24)
Now, we shall turn our attention to deriving an expression for the mean and variance of the difference:

\[
\hat{P}_l(\hat{\theta}) - \hat{P}_{l-1}(\hat{\theta}) = \log \left( \frac{|\Sigma_l(\hat{\theta})|}{|\Sigma_{l-1}(\hat{\theta})|} \right) \\
= \log \left( \frac{T + Ch^y_l}{T + Ch^y_{l-1}} \right) \\
= \log(1 + \frac{C}{T}h^y_l) - \log(1 + \frac{C}{T}h^y_{l-1}) \\
\approx \frac{C}{T} (h^y_l - h^y_{l-1})
\]

(4.25)

Therefore, \( \hat{P}_l(\hat{\theta}) - \hat{P}_{l-1}(\hat{\theta}) = \mathcal{O}(h^y_{l-1}) = \mathcal{O}(h^y_l) \). Thus, we can write:

\[
\mathbb{E}[\hat{P}_l(\hat{\theta}) - \hat{P}_{l-1}(\hat{\theta})] = \mathcal{O}(h^y_l).
\]

(4.26)

Since \( \text{Var}[\hat{P}_l(\hat{\theta}) - \hat{P}_{l-1}(\hat{\theta})] \leq \mathbb{E}[(\hat{P}_l(\hat{\theta}) - \hat{P}_{l-1}(\hat{\theta}))^2] \), then we obtain:

\[
\text{Var}[\hat{P}_l(\hat{\theta}) - \hat{P}_{l-1}(\hat{\theta})] = \mathcal{O}(h^{2y}_l).
\]

(4.27)

Since \( P_l(\hat{\theta}) = \hat{P}_l(\hat{\theta}) + \kappa \) where \( \kappa \) is independent from the discretization, then we can write:

\[
\mathbb{E}[P_l - P_{l-1}] = \mathcal{O}(h^y_l).
\]

\[
\text{Var}[P_l - P_{l-1}] = \mathcal{O}(h^{2y}_l).
\]

In the special case when \( h_l = h_0 2^{-l} \), we obtain:

\[
\mathbb{E}[P_l - P_{l-1}] = \mathcal{O}(2^{-\eta l}).
\]

\[
\text{Var}[P_l - P_{l-1}] = \mathcal{O}(2^{-2\eta l}).
\]

For a fixed number of repetitive experiment \( N_e \), the work of MLMC+LA method, \( W \), is mainly the work coming from the use of the MLMC method. Using the notations introduced in Giles theorem [40], we have \( \beta = 2\eta \) and \( \alpha = \eta \).
Assuming the cost per level is $O(2^{\gamma l})$, we can write the work of the MLMC+LA method as:

$$ W \propto \begin{cases} 
  TOL^{-2} & \text{if } \eta > \frac{\gamma}{2} \\
  TOL^{-2}(\log TOL)^2 & \text{if } \eta = \frac{\gamma}{2} \\
  TOL^{-2-(\gamma-2\eta)/\eta} & \text{if } \eta < \frac{\gamma}{2} 
\end{cases} $$

The error of the MLMC+LA method can be written as the sum of two errors:

$$ Error = \varepsilon_1 + \varepsilon_2, $$

where $\varepsilon_1$ is the error coming from the LA method and $\varepsilon_2$ is the MLMC error. In this case, combining the two methods would be advantageous if the errors coming from both methods are of the same magnitude, i.e. $\varepsilon_1 \simeq \varepsilon_2$. Recalling that $\varepsilon_2 \propto TOL$ and that $\varepsilon_1 \propto N_e^{-1}$, we can say that the work of the MLMC+LA method is given by:

$$ W \propto \begin{cases} 
  N_e^2 & \text{if } \eta > \frac{\gamma}{2} \\
  N_e^2(\log N_e)^2 & \text{if } \eta = \frac{\gamma}{2} \\
  N_e^{2+(\gamma-2\eta)/\eta} & \text{if } \eta < \frac{\gamma}{2} 
\end{cases} $$
Chapter 5

Numerical Results

In this chapter, we show the accuracy and efficiency of the aforementioned methods via two numerical examples.

5.1 Example 1

5.1.1 Problem set-up

In this first example, we look to estimate the expected information gain of the sensor deployment design for a Darcy flow problem governed by one-dimensional Poisson equation. We consider the following Dirichlet problem:

\[
\begin{cases}
-(\theta(x)u'(x))' = f(x), & x \in \Gamma \\
u(0) = u(1) = 0
\end{cases}
\] (5.1)

where \( f(x) = \sin(\pi x)\cos(\pi x) \) and \( \Gamma = ]0, 1[. \)

We look to determine the value of the expected information gain for a given experimental set-up. In this example, the pressure is measured using 13 sensors placed in the following locations \( \xi = [0.1, 0.15, 0.2, 0.25, 0.32, 0.4, 0.45, 0.55, 0.6, 0.68, 0.75, 0.81, 0.94] \).

To solve this forward problem, we use the Finite Element Method (FEM). The first
step consists in formulating the variational problem. For that, we define the following spaces:

\[ \mathcal{D}(\Gamma) := \{ \phi : \Gamma \to \Gamma : \phi \in C^\infty(\Gamma) \text{ and support}(\phi) \text{ is compact} \}, \]

\[ H^1_0(\Gamma) := \{ u \in L^2(\Gamma) : u' \in L^2(\Gamma) \text{ and } u(0) = u(1) = 0 \}. \]

where the gradient \( u' \) is defined in the weak sense, i.e. \( \int_\Gamma \phi u' = -\int_\Gamma u \phi', \forall \phi \in \mathcal{D}(\Gamma). \)

\( \mathcal{D}(\Gamma) \) is called the space of test functions and \( H^1_0(\Gamma) \) is a Hilbert space equipped with the norm \( \| u \|_{H^1_0(\Gamma)} := \| u' \|_{L^2(\Gamma)} = (\int_\Gamma |u'|^2)^{1/2} \).

Let \( u \) be solution of (5.1), then for every \( v \in \mathcal{D}(\Gamma) \), we have:

\[-\int_\Gamma v(\theta u') = \int_\Gamma f v. \tag{5.2}\]

Using Green’s formula, we get:

\[ \int_\Gamma \theta u' v' = \int_\Gamma f v. \tag{5.3}\]

We recall the Lax-Milgram theorem [45]:

**Theorem 3.** Let \( a \) be a bilinear, continuous and coercive form on a Hilbert space \( H \).

If \( l \) is a continuous linear form on \( H \), there is a unique \( u \in H \) such as:

\[ a(u, v) = l(v), \quad \forall u, v \in H \tag{5.4} \]

Let \( a(u, v) := \int_\Gamma \theta u' v' \) and \( l(v) := \int_\Gamma f v \), then we have:

\[ a(u, v) = l(v), \quad \forall u, v \in H^1_0(\Gamma) \tag{5.5} \]

We know that \( a \) is a bilinear form, \( l \) is a linear form and that \( H^1_0(\Gamma) \) is Hilbert space.
Using Cauchy-Schwartz inequality, we can write that $\forall u, v \in H^1_0(\Gamma)$:

$$|a(u, v)| \leq \kappa \|u\|_{L^2(\Gamma)} \|v\|_{L^2(\Gamma)} \leq \kappa \|u\|_{H^1_0(\Gamma)} \|v\|_{H^1_0(\Gamma)}, \quad (5.6)$$

where $\kappa$ is constant. This proves that $a$ is continuous on $H^1_0(\Gamma)$.

In addition to that, we have:

$$a(u, u) \geq C \|u\|^2_{H^1_0(\Gamma)}, \quad (5.7)$$

where $C$ is a constant.

Let $\Gamma = \Gamma_1 \cup \Gamma_2$ where $\Gamma_1 = [0, 1/2]$ and $\Gamma_2 = [1/2, 1]$ and we define $\theta$ as:

$$\theta(x) = \begin{cases} 
\theta_1 & \text{if } x \in \Gamma_1 \\
\theta_2 & \text{otherwise}
\end{cases}$$

where $\mathcal{U}(c, d)$ is the uniform distribution with bounds $c$ and $d$. Therefore, $\theta > 0$ and thus $C > 0$ which means combined with (5.7) that $a$ is coercive.

Using Theorem 3, we can state that the problem (5.5) has a unique solution.

Since $H^1_0(\Gamma)$ has infinite dimension, we will look to solve (5.5) in a finite dimensional space $V_h \subset H^1_0(\Gamma)$ with dimension $N_h$:

$$a(u_h, v_h) = l(v_h), \quad \forall u_h, v_h \in V_h \quad (5.8)$$

Let $(\phi_1, \phi_2, \ldots, \phi_{N_h})$ be a basis for the space $V_h$. In this case, we can write the solution as:

$$u_h(x) = \sum_{i=1}^{N_h} u_i \phi_i(x). \quad (5.9)$$

Since $V_h$ is a closed space of a Hilbert space, then it’s also a Hilbert space. Using Lax-Milgram theorem, the problem (5.8) has a unique solution $u_h$ given by (5.9) where
the coefficients \( \{ u_i \}_{i=1}^{N_h} \) are unique. We use piecewise linear continuous finite elements on a uniform grid \((\Gamma_h)_h\) on \([0,1]\).

The problem (5.8) is thus equivalent to finding \( U = (u_1, u_2, \ldots, u_{N_h}) \) solution of the linear system:

\[
A_h(\theta)U_h = F_h, \tag{5.10}
\]

where \((A_h)_{ij}(\theta) = a(\phi_j, \phi_i), \forall i, j = 1, \ldots, N_h\) and \((F_h)_i = l(\phi_i), \forall i = 1, \ldots, N_h\). The matrix \(A_h(\theta)\) is called stiffness matrix and the vector \(F_h\) is the force vector. To solve this problem, we use \(P_1\) finite elements type, see Appendix (A).

To have a look at the error estimate, we refer to the Céa’s lemma [46]:

**Lemma 4.** If \( u \) and \( u_h \) are respectively solutions of the problems (5.5) and (5.8), then we have the following inequality:

\[
\| u - u_h \| \leq \frac{\gamma}{\alpha} \| u - v \|, \quad \forall v \in V_h.
\]

Using this lemma, we can say that \( u_h \) is the best approximation of \( u \) in \( V_h \) up to a constant \( \frac{\gamma}{\alpha} \).

The model (2.1) can be written in the context of this problem as:

\[
y_i = P_h U_h(\theta_t) + \epsilon_i = P_h K_h(\theta_t)^{-1} F_h + \epsilon_i, \quad i = 1, \ldots, N_e \tag{5.11}
\]

where \( P_h \) is the interpolation matrix and \( \epsilon \sim \mathcal{N}(0, 10^{-7}I) \).

The \( k^{th} \) column of the Jacobian matrix of the model \( g \) with respect to the parameter \( \theta \) is given by:

\[
J_k = \frac{\partial g}{\partial \theta} = -P_h K_h^{-1} \sum_{\epsilon=1}^{NE} \frac{\partial K_e}{\partial \theta_k} K_h^{-1} F_h, \quad k = 1, \ldots, \text{dim}(\theta)
\]

where \( NE \) is the number of elements in the discretization and \( K_e \) is the element
stiffness matrix.

5.1.2 Results and discussions

Figures 5.1 and 5.2 show respectively the behavior of the logarithm base 2 of the variance and the mean of the quantities $P_l$ and $P_l - P_{l-1}$ as a function of the level $l$. We recall that in this work, $P_l$ represents the value of $D_{KL}$ computed on the level $l$. The dotted line is for the statistics of $P_l - P_{l-1}$ whereas the straight line represents the statistics of $P_l$. The red line represents the slopes $-2$ and $-1$ in the plots 5.1 and 5.2 respectively. The two figures are plotted using $10^5$ paths. For $l = 7$, we notice that $V_l$ is around 256 times smaller than the variance $V[P_l]$ of the standard MC method using the same time step. Also, for the same level, the relative error $\frac{\mathbb{E}[P_l - P_{l-1}]}{\mathbb{E}[P_l]}$ is of the order of $10^{-3}$. The MLMC convergence plots 5.1 and 5.2 confirm that numerical simulations match the theoretical order of convergence. In fact, by looking at the slope red line, we can state that:

$$\mathbb{E}[P_l - P_{l-1}] \approx \mathcal{O}(M_l^{-1}),$$
$$\text{Var}[P_l - P_{l-1}] \approx \mathcal{O}(M_l^{-2}).$$

Figures 5.3 and 5.4 show that the behavior of the number of sample per level $N_l$ with respect to the level $l$ for different tolerances $TOL$ when using the MLMC and the MLMC combined with the LA. Since $V_l = \mathcal{O}(2^{-\beta l})$ and $C_l = \mathcal{O}(2^{\gamma l})$, then using the expression of $N_l$ defined in Algorithm 3, we can see that $N_l = \mathcal{O}(2^{-\frac{1}{2}(\gamma+\beta)l})$. We can note that for a given $TOL$, $N_l$ is a decreasing function of $l$. This is due to the fact that $V_l$ also decreases with the level as shown in the plot 5.1. For a fixed level, the smaller the value of $TOL$, the greater the number of samples needed. Also, the value of the maximum level of $L$ increases as $TOL$ decreases. In fact, $L$ grows like $\log(TOL^{-1})$ since we require $h_L^\alpha = 2^{-\alpha L} = \mathcal{O}(TOL)$ in order to control the bias error.
Figure 5.1: $\log_2(\text{variance})$ as a function of the level $l$
The red line corresponds to the theoretical behavior, i.e. a slope of $-2$

Figure 5.2: $\log_2(|\text{mean}|)$ as a function of the level $l$
The red line corresponds to the theoretical behavior, i.e. a slope of $-1$
Figure 5.3: Number of samples per level for different tolerances using MLMC

The red line corresponds to the theoretical behavior, i.e. a slope of $-0.375$

In plot 5.5, we compare the computational cost of four methods as function of the accuracy $TOL$:

- The straight red line corresponds to the standard MC combined with the LA.
- The dashed red line corresponds to the MLMC combined with the LA.
- The straight blacked line corresponds to the DLMC.
- The dashed black line corresponds to the MLMC.
Figure 5.4: Number of samples per level for different tolerances using MLMC+LA
The red line corresponds to the theoretical behavior, i.e. a slope of $-0.6$.

Figure 5.5: Cost comparison between the different methods.
From this plot, we note that the MLMC method reduces significantly the computational cost. In fact, the dashed lines confirm that the cost of using MLMC is $O(TOL^{-2})$ which get along with the theory using the complexity Theorem of Giles. On the other hand, standard MC complexity is roughly of the order of $O(TOL^{-3})$. The ”staircase effect” in the straight red curve is due to the fact that $L = 7$ for $TOL = 10^{-2}$, $2 \times 10^{-2}$, $5 \times 10^{-2}$ and $L = 8$ for $TOL = 10^{-3}$, $2 \times 10^{-3}$ when using MLMC+LA, see figure 5.4. The LA provides a faster way to compute the expected information gain than the direct methods based on the DLMC. In table 5.1, we report the speed-up ratio of the MLMC+LA method compared to the other methods for $TOL = 10^{-2}$ and $TOL = 10^{-3}$. For instance, for $TOL = 10^{-3}$, the MLMC+LA method is 11 times faster than MLMC and MC+LA methods and around 570 times faster than the DLMC method. We can see that combining the LA with the MLMC or the standard MC gives better complexity results than just relying on the MLMC or the DLMC.

Table 5.1: Speed-up of the MLMC+LA compared to the other methods

<table>
<thead>
<tr>
<th>Method</th>
<th>$TOL = 10^{-2}$</th>
<th>$TOL = 10^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLMC</td>
<td>7.1</td>
<td>11.4</td>
</tr>
<tr>
<td>DLMC</td>
<td>34.2</td>
<td>571.4</td>
</tr>
<tr>
<td>MC+LA</td>
<td>5.7</td>
<td>11.4</td>
</tr>
</tbody>
</table>

In table 5.2, we report the computational errors of both the MLMC+LA as well as the MLMC methods for different tolerances.

Table 5.2: Computational errors of the MLMC+LA and the MLMC methods

<table>
<thead>
<tr>
<th>Method</th>
<th>$TOL = 10^{-2}$</th>
<th>$TOL = 5 \times 10^{-3}$</th>
<th>$TOL = 2 \times 10^{-3}$</th>
<th>$TOL = 10^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLMC</td>
<td>$8.4 \times 10^{-2}$</td>
<td>$3.8 \times 10^{-2}$</td>
<td>$3.2 \times 10^{-2}$</td>
<td>$2.5 \times 10^{-2}$</td>
</tr>
<tr>
<td>MLMC+LA</td>
<td>$2 \times 10^{-3}$</td>
<td>$10^{-3}$</td>
<td>$2.1 \times 10^{-4}$</td>
<td>$1.5 \times 10^{-4}$</td>
</tr>
</tbody>
</table>
5.2 Example 2

5.2.1 Notations

We start by introducing the notations used in this second example:

- $C_k$: volume concentration of component $k$ in the pore space
- $\phi$: porosity
- $\rho_k$: density of component $k$
- $C_{kl}$: volume concentration of component $k$ in phase $l$
- $S_l$: saturation of phase $l$
- $S_{lr}$: residual saturation of phase $l$
- $n_p$: total number of phases
- $u_l$: flux of phase $l$
- $J_{kl}$: flux due to diffusion
- $\mathbf{n}$: normal vector at the boundary surface
- $R_k$: source term which represents mass change due to sink, well or chemical reactions
- $P_l$: pressure of phase $l$
- $K$: absolute permeability
- $K_l$: relative permeability of phase $l$
- $K_l^0$: relative permeability end points of phase $l$
- $D_{kl}$: dispersion coefficient of component $k$ in phase $l$
5.2.2 Problem set-up

Enhanced Oil Recovery (EOR) are methods attempting to improve the extraction of crude oil by injecting different types of materials that can modify chosen properties of fluids (water and oil) in the reservoir [47, 48]. In fact, 20% to 40% of the reservoir original oil can be extracted by conventional recovery techniques [49, 50]. However, EOR methods enable to capture an additional 10% to 40% of the Original Oil In Place (OOIP)[51].

EOR techniques can be divided into three main classes: thermal, miscible and chemical methods. Surfactant polymer flooding is a chemical EOR method.

The use of surfactants in the oil industry goes back to the year 1980. Once injected, surfactants can reduce capillary forces by creating low interfacial tensions,
between the hydrocarbon and the injected fluid, which leads to the movement of the trapped oil. In EOR, surfactant can be combined with a polymer flooding. The use of polymer helps creating microemulsions that aim to improve the sweep efficiency of the reservoir.

5.2.3 Mathematical formulation of the forward problem

In this section, we present briefly the main equations that characterize the forward model.

The mass conservation equation for the $k^{th}$ component is given by:

$$
\int_V \frac{\partial}{\partial t} (\phi \rho_k C_k) dV = -\int_V \sum_{l=1}^{n_p} \rho_k (C_{kl} u_l + J_{kl}) \cdot n \, d\Gamma + \int_V R_k dV.
$$

(5.12)

Using the divergence theorem, we get:

$$
\int_V \frac{\partial}{\partial t} (\phi \rho_k C_k) dV = -\int_V \nabla \cdot \left[ \sum_{l=1}^{n_p} \rho_k (C_{kl} u_l + J_{kl}) \right] dV + \int_V R_k dV.
$$

(5.13)

Since the equation is written for an arbitrary volume in the porous media, we deduce that:

$$
\frac{\partial}{\partial t} (\phi \rho_k C_k) + \nabla \cdot \left[ \sum_{l=1}^{n_p} \rho_k (C_{kl} u_l + J_{kl}) \right] = R_k.
$$

(5.14)

The flux is related to the gradient of concentration by means of Fick’s law of diffusion:

$$
J_{kl} = -\phi S_l D_{kl} \nabla C_{kl}.
$$

(5.15)

Replacing (5.15) in (5.14), we get:

$$
\frac{\partial}{\partial t} (\phi \rho_k C_k) + \nabla \cdot \left[ \sum_{l=1}^{n_p} \rho_k (C_{kl} u_l - \phi S_l D_{kl} \nabla C_{kl}) \right] = R_k.
$$

(5.16)
Darcy’s law of phase flux gives the expression of the Darcy velocity \( u_l \) as:
\[
  u_l = -\frac{K K_l}{\mu_l} \nabla P_l.
\]

Combining equations (5.16), (5.17) and the definition of capillary pressure, we get the pressure equation for phase \( l' \):
\[
  \phi c_l \frac{\partial P_{l'}}{\partial t} + \nabla \cdot K \cdot \lambda_{rTc} \nabla P_{l'} = -\nabla \cdot \sum_{l=1}^{n_p} K \cdot \lambda_{rTc} \nabla h + \nabla \cdot \sum_{l=1}^{n_p} K \cdot \lambda_{rTc} \nabla P_{cll'} + \sum_{k=1}^{n_{cv}} Q_k.
\]

The total volume of the \( k^{th} \) component per unit volume can be expressed as:
\[
  \tilde{C}_k = (1 - \sum_{k=1}^{n_{cv}} \hat{C}_k) C_k + \hat{C}_k.
\]

The system of differential equations to be solved consists of equations (5.18) and (5.19). Finite difference schemes are used to solve this system. The temporal discretization is implicit in pressure and explicit in concentration. We assume that there is no convection, no dispersion and no thermal flux through all the boundaries.

For the forward model solution, we use the University of Texas CHEMical compositional simulator (UTCHEM) [52, 53, 54, 55, 56, 57]. UTCHEM is a three-dimensional multiphase, multicomponent chemical flooding simulator created by the Center for Petroleum and Geosystems of Texas University. In this work, we assume an homogeneous permeability and the existence of three phases: water (w), oil (o) and microemulsion (m). The horizontal core is flooded with water, surfactant and polymer.

The relative permeability curves are defined in our model using Corey type functions. For the phase \( l \), the relative permeability is given by:
\[
  K_l = K_l^0 \left( \frac{S_l - S_{lr}}{1 - \sum_{j=1}^{n_p} S_{lj}} \right)^{n_l} \quad \forall l = 1, \ldots, n_p
\]
where $n_l$ is the Corey exponent of phase $l$.

The parameters of the model are:

- Endpoint relative permeability of the different phases ($K_w^0, K_o^0, K_m^0$)
- Residual saturation ($S_{rw}, S_{ro}, S_{rm}$)
- Relative permeability exponents ($n_o, n_w, n_m$)

Thus, we can define the parameter as $\theta = [K_w^0, K_o^0, K_m^0, S_{rw}, S_{ro}, S_{rm}, n_o, n_w, n_m]^T$.

The control variables are:

- Injection rate ($r$)
- Volume fraction of injected surfactant ($c_s$)
- Weight concentration of injected polymer ($c_p$)
- Total volume injected $V$

Thus, the experimental set-up is $\xi = [r, c_s, c_p, V]^T$. In our study, we will modify the experimental set-up by changing the injection rate and maintaining the other control variables fixed.

In this work, we model the parameter $\theta$ as a random vector where each component has a uniform pdf. Table 5.3 details the bounds of the pdf for each the relative permeability parameters.

<table>
<thead>
<tr>
<th>Phase</th>
<th>$S_{rX}$</th>
<th>$K_X$</th>
<th>$n_X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>[0.25, 0.35]</td>
<td>[0.25, 0.35]</td>
<td>[2, 4]</td>
</tr>
<tr>
<td>Oil</td>
<td>[0.28, 0.38]</td>
<td>[0.55, 0.65]</td>
<td>[2, 4]</td>
</tr>
<tr>
<td>Microemulsion</td>
<td>[0.25, 0.35]</td>
<td>[0.25, 0.35]</td>
<td>[2, 4]</td>
</tr>
</tbody>
</table>

In the core flooding example discussed in this section, we assume that we have constant porosity for the whole reservoir and that there is no water nor air tracers.
The reservoir contains six components: water, oil, surfactant, polymer, anion and cation. The permeability is constant along the x-axis. Two wells are used in this example. The output we are interested in is the cumulative oil recovery of the second well.

Table 5.4 summarizes the main simulation data used in this example and figure 5.6 presents a schematic view of the studied porous media.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porous media dimension</td>
<td>$0.745 \times 0.11 \times 0.11 \text{ ft}^3$</td>
</tr>
<tr>
<td>Number of grid blocks</td>
<td>$80 \times 1 \times 1$</td>
</tr>
<tr>
<td>Porosity</td>
<td>0.219 pore volume</td>
</tr>
<tr>
<td>Permeability</td>
<td>72 millidarcy</td>
</tr>
<tr>
<td>Initial water saturation</td>
<td>0.59 pore volume</td>
</tr>
<tr>
<td>Water viscosity</td>
<td>0.678 centipose</td>
</tr>
<tr>
<td>Oil viscosity</td>
<td>7 centipose</td>
</tr>
<tr>
<td>Specific weight of water</td>
<td>0.433 psi/ft</td>
</tr>
<tr>
<td>Specific weight of oil</td>
<td>0.377 psi/ft</td>
</tr>
<tr>
<td>Initial injection period</td>
<td>2.5 pore volume</td>
</tr>
</tbody>
</table>

There are three sequences of injection: a first volume that includes surfactant and polymer, a second slug that only contains polymer and a final one consists of water. Table 5.5 details the injection strategy with regards to concentrations and volume of injection in each slug.

<table>
<thead>
<tr>
<th>Slug number</th>
<th>Water</th>
<th>Oil</th>
<th>Surfactant</th>
<th>Polymer</th>
<th>Anion</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>0.99</td>
<td>0</td>
<td>0.01</td>
<td>0.15</td>
<td>0.3</td>
</tr>
<tr>
<td>Second</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.15</td>
<td>0.26</td>
</tr>
<tr>
<td>Third</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.26</td>
</tr>
</tbody>
</table>

The first three columns of this table represent the amount of water, oil and surfactant injected in volume fraction. The sum of the volume fractions must add up to one. The fourth column is the amount of polymer injected in % of
concentration. The last column is the amount of anion concentration in meq/ml which is an indicator of water salinity.

Figure 5.6: Schematic view of the porous media

5.2.4 Results and discussion

We start by studying the convergence of the statistics, mean and variance, of the quantity $P_l$ as well as the difference $P_l - P_{l-1}$ as a function of the level $l$. The results are plotted in figures 5.7 and 5.8. We remark, for level $l = 7$, the variance of the difference is around 350 times smaller than the variance of the standard MC. The relative error for this level is approximately $8 \times 10^{-3}$. In both plots, we compare the numerical results to the theoretical order of convergence given by the red lines of slopes $-2$ and $-1$ for the variance and the mean respectively. We verify that the numerical simulations are in harmony with the theoretical results obtained in section
4.2.2 of chapter 4.

Figure 5.7: $\log_2(\text{variance})$ as a function of the level $l$
The red line corresponds to the theoretical behavior, i.e. a slope of $-2$

Figure 5.8: $\log_2 |\text{mean}|$ as a function of the level $l$
The red line corresponds to the theoretical behavior, i.e. a slope of $-1$
For different tolerances $TOL$, we plot the number of sample per level $N_l$ as a function of the level $l$ when using the MLMC in figure 5.9 and when using the MLMC combined with the LA in figure 5.10. The optimal number of sample per level $N_l$ behaves theoretically as $\mathcal{O}(2^{-\frac{1}{2}(\gamma+\beta)l})$. Therefore, $N_l$ is a decreasing function of $l$ for a fixed value of $TOL$ as shown in both plots 5.9 and 5.10. For a given level, the number of samples needed per level increases as the tolerance $TOL$ becomes smaller. Also, the value of the maximum level of $L$ grows as $\log(TOL^{-1})$ and thus $L$ increases as $TOL$ decreases.

In plot 5.11, we compare the computational cost of the four methods, MLMC, DLMC, MC+LA and MLMC+LA, as function of the accuracy $TOL$.

![Figure 5.9: Number of samples per level for different tolerances using MLMC](image)

The red line corresponds to the theoretical behavior, i.e. a slope of $-0.375$
Figure 5.10: Number of samples per level for different tolerances using MLMC+LA
The red line corresponds to the theoretical behavior, i.e. a slope of $-0.6$

Figure 5.11: Cost comparison between the different methods
From plot 5.11, we can see that the computational cost is reduced when the MLMC method is used. In agreement with the complexity Theorem of Giles, the dashed lines indicate that the cost of using MLMC is $O(TOL^{-2})$. Standard MC complexity is roughly of the order of $O(TOL^{-3})$ which can be seen from the straight lines behavior. The estimation of the expected information gain is accelerated when using the LA. In table 5.6, we report the speed-up ratio of the MLMC+LA method compared to the other methods for $TOL = 10^{-2}$ and $TOL = 2 \times 10^{-3}$. For instance, for $TOL = 2 \times 10^{-3}$, the MLMC+LA method is 110 times faster than MLMC and around 2040 times faster than the DLMC method. We can see that combining the LA with the MLMC method gives the best complexity compared to the other methods.

Table 5.6: Speed-up of the MLMC+LA compared to the other methods

<table>
<thead>
<tr>
<th>Method</th>
<th>$TOL = 10^{-2}$</th>
<th>$TOL = 2 \times 10^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLMC</td>
<td>110.8</td>
<td>110.4</td>
</tr>
<tr>
<td>DLMC</td>
<td>1166.9</td>
<td>2041.1</td>
</tr>
<tr>
<td>MC+LA</td>
<td>111.2</td>
<td>1576.9</td>
</tr>
</tbody>
</table>
Chapter 6

Conclusion

6.1 Summary

In this work, we presented numerous methods (MLMC, DLMC, MLMC+LA and MC+LA) for the estimation of the expected information gain. The latter quantity is a measure often used as a utility function to evaluate the goodness of experimental designs in the Bayesian framework.

We tested these methods on two numerical examples: the first example is the design of sensor deployment for a Darcy flow problem governed by a one-dimensional Poisson equation and the second is the study of a chemical Enhanced Oil Recovery (EOR) core flooding experiment assuming homogeneous permeability.

A comparison between the different methods shows that the LA combined with the MLMC method turns to be the most efficient method in terms of computational cost. We also provide an alternative approach, the direct MLMC, when the LA fails and which is faster than the standard DLMC.

6.2 Future Work

In the light of this thesis, the conducted work can be guidance for future research. In fact, many possible extensions can be done. We can extend this work by using
stochastic optimization algorithm to determine the optimal experimental set-up for a given experiment. Another possible extension is to consider more complex noise model, for instance error in measurement model. In this model, the experimental set-up is modeled as a random variable subject to uncertainty rather than a deterministic set of variables. In this case, we can consider using Multi-index Monte Carlo, [58], to define the levels in order to further improve the computational cost.
REFERENCES


A Nested MLMC estimator

In this appendix, we derive the expression of the nested MLMC estimator. First, we recall the expression of the estimator:

\[
Y_l = \frac{1}{N_l} \sum_{n=1}^{N_l} \left[ \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M_l} \sum_{m=1}^{M_l} p(y_n|\theta_m)} \right) - \frac{1}{2} \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M_{l-1}} \sum_{m=1}^{M_{l-1}} p(y_n|\theta_m)} \right) \right] - \frac{1}{2} \log \left( \frac{p(y_n|\theta_n)}{\frac{1}{M_{l-1}} \sum_{m=M_{l-1}+1}^{M_l} p(y_n|\theta_m)} \right).
\]

(A.1)

We start by defining the following quantities:

\[
\frac{1}{M_{l-1}} \sum_{m=1}^{M_{l-1}} \psi(y_n, \theta_m) = \mathbb{E}[\psi(y_n, \theta_m)] + \Delta \psi_1^{(n)}.
\]

\[
\frac{1}{M_{l-1}} \sum_{m=M_{l-1}+1}^{M_l} \psi(y_n, \theta_m) = \mathbb{E}[\psi(y_n, \theta_m)] + \Delta \psi_2^{(n)}.
\]

Since \( \frac{1}{M_l} = \frac{1}{2} \times \frac{1}{M_{l-1}} \), then we can write:

\[
\frac{1}{M_l} \sum_{m=1}^{M_l} \psi(y_n, \theta_m) = \mathbb{E}[\psi(y_n, \theta_m)] + \frac{1}{2} (\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)}).
\]

The second order Taylor expansions of each term in (A.1) are given by:
Thus, we get the expression of the nested MLMC given in (A.1).

Going back to (A.1) and replacing the Taylor expansions, we get:

\[ f \left( \frac{1}{M_i} \sum_{m=1}^{M_i} \psi(y_n, \theta_m) \right) = f \left( E[\psi(y_n, \theta_m)] + \frac{1}{2}(\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)}) \right) \]

\[ = f(E[\psi(y_n, \theta_m)] + \frac{1}{2}(\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)})) + \frac{1}{8}(\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)})^2 \]

\[ + O((\Delta \psi_1^{(n)}) + \Delta \psi_2^{(n)})^3). \]

\[ f \left( \frac{1}{M_{t-1}} \sum_{m=1}^{M_{t-1}} \psi(y_n, \theta_m) \right) = f \left( E[\psi(y_n, \theta_m)] + \Delta \psi_1^{(n)} \right) \]

\[ = f(E[\psi(y_n, \theta_m)]) + \Delta \psi_1^{(n)} f'(E[\psi(y_n, \theta_m)]) + \frac{1}{2}(\Delta \psi_1^{(n)})^2 f''(E[\psi(y_n, \theta_m)]) + O((\Delta \psi_1^{(n)})^3). \]

\[ f \left( \frac{1}{M_{t-1}} \sum_{m=M_{t-1}+1}^{M_t} \psi(y_n, \theta_m) \right) = f \left( E[\psi(y_n, \theta_m)] + \Delta \psi_2^{(n)} \right) \]

\[ = f(E[\psi(y_n, \theta_m)]) + \Delta \psi_2^{(n)} f'(E[\psi(y_n, \theta_m)]) + \frac{1}{2}(\Delta \psi_2^{(n)})^2 f''(E[\psi(y_n, \theta_m)]) + O((\Delta \psi_2^{(n)})^3). \]

Going back to (A.1) and replacing the Taylor expansions, we get:

\[ Y_i = \frac{1}{N_i} \sum_{l=1}^{N_i} f(E[\psi(y_n, \theta_m)]) + \frac{1}{2}(\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)}) f'(E[\psi(y_n, \theta_m)]) + \frac{1}{8}(\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)})^2 \]

\[ f''(E[\psi(y_n, \theta_m)]) + O((\Delta \psi_1^{(n)} + \Delta \psi_2^{(n)})^3) - \frac{1}{2} f'(E[\psi(y_n, \theta_m)]) \]

\[ - \frac{1}{4}(\Delta \psi_1^{(n)})^2 f''(E[\psi(y_n, \theta_m)]) + O((\Delta \psi_1^{(n)})^3) - \frac{1}{2} f'(E[\psi(y_n, \theta_m)]) \]

\[ - \frac{1}{4}(\Delta \psi_2^{(n)})^2 f''(E[\psi(y_n, \theta_m)]) + O((\Delta \psi_2^{(n)})^3) \]

\[ \approx \frac{1}{N_i} \sum_{n=1}^{N_i} \left[ - \frac{1}{8}(\Delta \psi_1^{(n)})^2 - \frac{1}{8}(\Delta \psi_2^{(n)})^2 + \frac{1}{4} \Delta \psi_1^{(n)} \Delta \psi_2^{(n)} \right] f''(E[\psi(y_n, \theta_m)]) \]

\[ \approx - \frac{1}{8} N_i \sum_{n=1}^{N_i} \left[ (\Delta \psi_1^{(n)})^2 + (\Delta \psi_2^{(n)})^2 - 2 \Delta \psi_1^{(n)} \Delta \psi_2^{(n)} \right] f''(E[\psi(y_n, \theta_m)]). \]

Thus, we get the expression of the nested MLMC given in (A.1).
B Finite element method in 1D

In this appendix, we explain further the finite element method (FEM) spirit through a 1D example. For simplicity, we assume that $\theta$ is a fixed value equal to 1. In this case, the problem writes:

\[
\begin{aligned}
-u''(x) &= f(x), \quad x \in [0, 1] \\
u(0) &= u(1) = 0
\end{aligned}
\]

The linear system is given by: $A_h U_h = F_h$ where $a(u, v) := \int_0^1 u'v'$ and $l(v) := \int_0^1 fv$.

We consider a uniform grid on $[0, 1]$: $x_i = ih$, $i = 0, \ldots, N_h$ where $h = \frac{1}{N_h+1}$ and we define the space, $\mathbb{P}_k$, of polynomials of degree less or equal to $k$:

$$
\mathbb{P}_k = \{ p(x) = \sum_{j=0}^{k} a_j x^j, a_j \in \mathbb{R} \}
$$

In this appendix, we shall explain briefly the finite elements approximation using $\mathbb{P}_1$ elements. We introduce the finite dimension functional space:

$$
V_h = \{ v \in C^0([0, 1]), v_{|[x_i,x_{i+1}]} \in \mathbb{P}_1, 0 \leq i \leq N_h, v(0) = v(1) = 0 \}
$$

A basis of $V_h$ is given by the set of continuous functions, $(\phi_1, \phi_2, \ldots, \phi_{N_h})$, linear on each element $[x_i, x_{i+1}]$ and such that:

$$
\forall i = 1, \ldots, N_h, \quad \forall j = 1, \ldots, N_h + 1, \quad \phi_i(x_j) = \delta_{i,j}
$$
where $\delta_{i,j}$ is the de Kronecker symbol.

We introduce the following function:

$$\phi(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

thus, every basis function $\phi_i$ has the following expression $\phi_i(x) = \phi\left(\frac{x - x_i}{h}\right)$.

In the FEM, each basis function $\phi_i$ has a very reduced support, that is to say that the set of all $x$ such that $\phi_i(x) \neq 0$ is small compared to the domain $[0, 1]$. This has the result that most of the coefficients of the stiffness matrix $A_h$ are null. In fact, for the problem (B.1), we have:

$$(A_h)_{i,j} = a(\phi_i, \phi_j) = \int_0^1 \phi'_i(x)\phi'_j(x)dx. \quad (B.1)$$

For $j$ different from $i$, $i-1$ and $i+1$, $(A_h)_{i,j} = 0$ since the functions $\phi_i$ and $\phi_j$ haven’t disjoint supports. The non-null elements can be computed easily:

$$(A_h)_{i,i-1} = a(\phi_i, \phi_{i-1}) = -\frac{1}{h},$$

$$(A_h)_{i,i-1} = a(\phi_i, \phi_i) = \frac{2}{h},$$

$$(A_h)_{i,i+1} = a(\phi_i, \phi_{i+1}) = -\frac{1}{h}.$$
To compute the vector $F$, the following integrals need to be computed:

$$(F_h)_i = l(\phi_i) = \int_0^1 f(x)\phi_i(x)dx = \int_{x_{i-1}}^{x_i} f(x)\phi_i(x)dx + \int_{x_i}^{x_{i+1}} f(x)\phi_i(x)dx.$$ 

In general, this integral cannot be calculated accurately because the function $f$ can be complicated. In practice, we use numerical integration techniques where, on each interval $[x_i, x_{i+1}]$, the integral approach by a quadrature formula. In our case, we use the midpoint quadrature:

$$\int_{x_i}^{x_{i+1}} h(x)dx = (x_{i+1} - x_i) \ h \left( \frac{x_i + x_{i+1}}{2} \right) + O((x_{i+1} - x_i)^2).$$

At this point, all is left is to take into consideration the boundary conditions and incorporates them in the linear system. Finding the vector $U$ gives us the function $u_h$ which converges to $u$ as $h \to 0.$
C Proof of the approximation of $\hat{\theta}$

In this appendix, we look to prove that when $N_e \to +\infty$, we can show that:

$$\hat{\theta} = \theta_t + \mathcal{O}_P\left(\frac{1}{\sqrt{N_e}}\right) \quad \text{(C.1)}$$

We recall that:

$$\hat{\theta} = \theta_t - (N_e J^T \Sigma_{\epsilon}^{-1} J + H^T \Sigma_{\epsilon}^{-1} E_S - \nabla \nabla h(\theta_t))^{-1} J^T \Sigma_{\epsilon}^{-1} E_S + \mathcal{O}_P\left(\frac{1}{N_e}\right). \quad \text{(C.2)}$$

So basically, we need to show that:

$$(N_e J^T \Sigma_{\epsilon}^{-1} J + H^T \Sigma_{\epsilon}^{-1} E_S - \nabla \nabla h(\theta_t))^{-1} J^T \Sigma_{\epsilon}^{-1} E_S = \mathcal{O}_P\left(\frac{1}{\sqrt{N_e}}\right). \quad \text{(C.3)}$$

In this proof, we use Einstein notation for the sum and we consider that the variance of the $k^{th}$ residual $r_i$ is $\text{Var}[r_{i,k}] = \sigma^2$. Since $\{r_i\}_{i=1}^{N_e}$ are i.i.d, then $\text{Var}[E_{S,k}] = N_e \sigma^2$. We have:

$$\mathbb{E}[J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk} E_{S,k}] = J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk} \mathbb{E}[E_{S,k}] = 0, \quad \text{(C.4)}$$

$$\text{Var}[J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk} E_{S,k}] = (J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk})^2 \text{Var}[E_{S,k}] = N_e (J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk})^2 \sigma^2.$$

Using Chebyshev’s inequality, we can write, $\forall \kappa > 0$, that:

$$\mathbb{P}(|J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk} E_{S,k}| > \kappa \sigma (J^T_{ij}(\Sigma_{\epsilon}^{-1})_{jk}) \sqrt{N_e}) < \frac{1}{\kappa^2}. \quad \text{(C.5)}$$
Let $\alpha = \frac{1}{\kappa^2}$ and $\eta = \kappa \sigma (\mathbf{J}^T_d (\Sigma^{-1}_e)_{jk})$, then we can write:

$$\mathbb{P}(|\mathbf{J}^T_d (\Sigma^{-1}_e)_{jk} \mathbf{E}_{S,k}| > \eta \sqrt{N_e}) < \alpha. \quad (C.6)$$

In other terms:

$$\mathbb{P}(|\mathbf{J}^T_d (\Sigma^{-1}_e)_{jk} \mathbf{E}_{S,k}| \leq \eta \sqrt{N_e}) \geq 1 - \alpha. \quad (C.7)$$

Since the last inequality holds for arbitrary $i$, $j$ and $k$, then we obtain:

$$\mathbb{P}(|\mathbf{J}^T \Sigma^{-1}_e \mathbf{E}_S| \leq \eta \sqrt{N_e}) \geq 1 - \alpha. \quad (C.8)$$

Therefore:

$$\mathbf{J}^T \Sigma^{-1}_e \mathbf{E}_S = O_P(\sqrt{N_e}).$$

Using a similar reasoning, we can show that:

$$\mathbf{H}^T \Sigma^{-1}_e \mathbf{E}_S = O_P(\sqrt{N_e}).$$

We easily can state that:

$$N_e \mathbf{J}^T \Sigma^{-1}_e \mathbf{J} - \nabla \nabla h(\theta_t) = O_P(N_e).$$

Thus:

$$\left( N_e \mathbf{J}^T \Sigma^{-1}_e \mathbf{J} + \mathbf{H}^T \Sigma^{-1}_e \mathbf{E}_S - \nabla \nabla h(\theta_t) \right)^{-1} \mathbf{J}^T \Sigma^{-1}_e \mathbf{E}_S = O_P\left(\frac{\sqrt{N_e}}{N_e}\right) = O_P\left(\frac{1}{\sqrt{N_e}}\right).$$
D Equivalence between difference of entropy and expected information gain

In this appendix, we provide the proof for the equivalence between difference of entropy and expected information gain. We will proof the following equality:

\[ I = -\Delta h = h(\theta) - h(\theta|y), \]  

where \( h(\theta) = -\int_\Theta p(\theta) \log(p(\theta)) d\theta. \)

We start by recalling the expression of the expected information gain:

\[ I = \int_\Theta \int_Y p(\theta,z) \log\left(\frac{p(\theta|z)}{p(\theta)}\right) p(z) d\theta dz. \]  

Thus:

\[
\begin{align*}
I &= \int_Y \int_\Theta p(\theta,z) \log\left(\frac{p(\theta,z)}{p(\theta)p(z)}\right) d\theta dz \\
&= \int_Y \int_\Theta p(\theta,z) \log(p(\theta,z)) d\theta dz - \int_Y \int_\Theta p(\theta,z) \log(p(\theta)) d\theta dz \\
&= \int_Y p(z) \left[ \int_\Theta p(\theta|z) \log(p(\theta|z)) d\theta \right] dz - \int_\Theta \log(p(\theta)) \left[ \int_Y p(\theta,z) d\theta \right] d\theta \quad \text{(D.3)} \\
&= -\int_Y p(z) h(\theta|y = z, \xi) dz - \int_\Theta p(\theta) \log(p(\theta)) d\theta \\
&= h(\theta) - h(\theta|y).
\]