Multilevel Monte Carlo for Stochastic PDEs

Inti-Raymi Carhuancho Mantripp
Mansfield College
University of Oxford

A dissertation submitted for the degree of
MSc in Mathematical Modelling & Scientific Computing
Trinity 2025

Contents

List of Figures

Chapter 1

Introduction

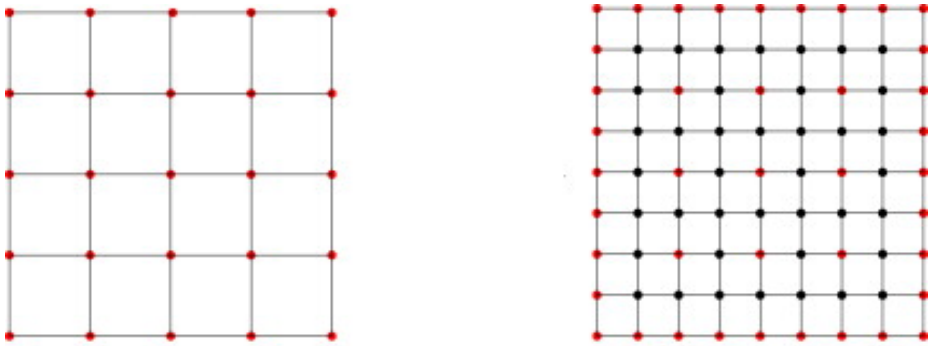
The purpose of this dissertation is to investigate the application of the Multilevel Monte Carlo (MLMC) method as a technique for reducing the computational cost in estimating expectations of random variables for two classes of parabolic stochastic partial differential equations (SPDEs).

The motivation for this stems from the fact that the standard Monte Carlo (MC) estimator suffers from slow convergence. Its root mean square error (defined in section SECTION HERE) decays $O(N^{-1/2})$ where N is the number of independent samples obtained, and therefore achieving a high accuracy estimate requires a high number of samples. As most SPDEs do not admit analytic solutions, the main approach available for estimating expectations of functionals of SPDEs is Monte Carlo methods. In the context of SPDEs, obtaining samples typically requires discretising the spatial domain via finite difference or finite element methods and then evolving the system to obtain the desires quantity. Accurate estimates requires fine discretisations, and as such, the cost of obtaining a sample can become expensive. This, combined with the slow convergence of the standard MC estimator, can make the standard Monte Carlo method prohibitively expensive.

A number of techniques exist that aim to improve on the convergence of the MC estimator. Their goal is to achieve an equivalent variance in their estimate at a reduced computational cost. The MLMC method is one such method. At its heart is an approach of computing a blend of cheaper and more expensive samples, unlike the MC estimator which uses only samples of the same cost. Many inexpensive samples provide the bulk of the statistical signal, while comparitively few more costly samples then correct any residual bias. In the context of SPDEs, this means obtaining samples using meshes at varying *levels* of discretisation. By taking many samples at cheaper, coarser levels and fewer at finer, more expensive levels, the MLMC estimator aim

to achieve an estimate of equivalent accuracy at reduced cost compared to the MC estimator.

The practical efficiency of the MLMC method however relies on correlating samples obtained at different levels, in a manner analogous to that of the control variates method [8] MORE CITATIONS HERE. This is because the cost savings is dependent on the extent to which samples across two adjacent levels of accuracy covary. Practically, this entails aligning the randomness used to generate pairs of coarse and fine samples. This is referred to as *coupling*. To provide intuition as to what this means, Figure 1.1 illustrates what the finite difference grid used by two adjacent estimators might look like. At each node in these grids, a random noise is required to be generated. This raises a serious question: what methods are best suited to aligning the randomness across samples? Further, how large an impact can different coupling methods have?



Coarse Grid Fine Grid

Figure 1.1: An illustrative coarse and fine finite difference grid used to obtain samples of $P_{\ell-1}$ and P_ℓ respectively. Red dots indicate vertices of the coarse grid. These are overlayed on the fine grid also, indicating common points both grids share and that both would evaluate. Black dots indicate vertices of the fine grid that are not common to the coarse grid.

There are three main questions this dissertation aims to answer therefore.

1. **What cost savings can MLMC achieve over standard MC in the context of two parabolic SPDEs?**

2. **What are the effects of different coupling mechanisms?**

3. **Monte Carlo methods are embarrasingly parallelisable. What further cost reductions are therefore achievable in high performance implementations of these methods if we take advantage of the highly parallel nature of MC methods?**

To answer these questions, this dissertation is structured as follows.

SECTION HERE provides the theoretical background and a review of the relevant literature. We begin by formally defining the standard MoION HERE To best present the answers found to these questions, this dissertation is structured as so. In the next section we formally present the requisite topics as they pertain to this investigation. This will cover the main theory of MLMC, a sufficiently relevant overview of SPDEs and a literature review of the current landscape.

SECTION HERE details our numerical implementation. We describe the finite difference schemes used to discretise the SPDEs and, crucially, the specific coupling strategies employed to correlate samples between levels. We then validate our implementation, demonstrating that it achieves the correct theoretical convergence rates for both the stochastic heat equation and the Dean-Kawasaki equation.

Section SECTION HERE will then present an MLMC implementation for the stochastic heat equation. We validate this implementation through a series of methods, varying the quantity of interest being derived, that our MLMC implementation converges to this and that its rate of convergence is appropriate. We then further validate our MLMC implementation of the Dean-Kawasaki equation, replicating the results of the original paper and our implementation.

Section SECTION HERE presents a comprehensive performance analysis. We directly address our first two research questions by comparing the computational cost of our MLMC implementation against a standard MC estimator for a range of target accuracies. Furthermore, we analyse the practical impact of different coupling mechanisms on overall efficiency.

Section SECTION HERE explores the potential for further cost reductions through high-performance computing. We discuss the design of a parallelised MLMC algorithm that leverages modern multi-core architectures and present performance metrics, such as computational speedup, to quantify the benefits of this approach.

Finally, Section SECTION HERE Finally, Chapter 6 concludes the dissertation. We summarise our key findings, revisit the initial research questions, and discuss potential avenues for future investigation.

Chapter 2

Preliminary Concepts and Methods

This section presents the essential thoeretical groundwork upon which the investigations of this dissertation is built. First, we present and develop the theory of the Monte Carlo (MC) and Multilevel Monte Carlo (MLMC) methods, analysing their respective convergence and computational complexities.

Second, we provide a concise introduction to the theory of stochastic partial differential equations (SPDEs). We focus on two parabolic equations which serve as our case studies: the Stochastic Heat Equation and the Dean-Kawasaki equation. We outline their mathematical formulations and briefly discuss the challenges associated with their numerical solutions, this being particularly pertinent to the Dean-Kawasaki equation.

Finally, we conclude the section with a literature rveiew. This review surveys the development and application of the MLMC method within the context of SPDEs, thereby situating the specific questions of this dissertation within the broader research performed in this area.

2.1 Monte Carlo and Multilevel Monte Carlo

The primary objective in many applications involving SPDEs is not to find a single, particular solution, but rather to compute the expected value of a quantity of interest that depends on the solution. For example, we may want to find the average temporature at a specific point in a domain governed by the stochastic heat equation, or the average particle density in a system described by the Dean-Kawasaki equation.

Let P represent such a quantity of interest, which we assume to be a real-valued random variable. Our goal throughout this section is to develop an efficient numerical

method for estimating its expectation, $\mathbb{E}[P]$. Since analytical expressions for $\mathbb{E}[P]$ are rarely available in this context, we must turn to computational methods. The most fundamental of these is defined below.

Definition 1 (Monte Carlo Estimator) *Let $\{P^{(n)}\}$ for $n = 1, \ldots, N$ be a set of N independent and identically distributed samples of a random variable P. The standard Monte Carlo estimator, $\hat{P}_{MC}$, of the expectation $\mathbb{E}[P]$ is the sample mean:*

$$\hat{P}_{MC} = \frac{1}{N} \sum_{n=1}^{N} P^{(n)}$$

By linearity of the expectation operator, the MC estimator in Definition 1 is unbiased. The accuracy of an estimator is typically estimated via its Mean Squared Error (MSE) and Root Mean Square Error (RMSE), which we define now.

Definition 2 (Mean Squared Error and Root Mean Squared Error) *Let P be a fixed unknown quantity and let $\hat{P}$ be an estimator for P. The **Mean Square Error (MSE)** of the estimator is the expected value of the squared error:*

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

This can be decomposed into terms representing the estimator's variance and squared bias:

$$MSE(\hat{P}) = \underbrace{\mathbb{V}[\hat{P}]}_{Variance\ of\ estimator} + \underbrace{(\mathbb{E}[\hat{P}] - \mathbb{E}[P])^2}_{Bias^2}$$

*The **Root Mean Square Error (RMSE)** is the square root of the MSE.*

Since the standard MC estimator is unbiased, its bias term is zero. Its MSE is therefore equal to its variance:

$$\text{MSE}(\hat{P}_{MC}) = \mathbb{V}[\hat{P}_{MC}] = \frac{1}{N}\mathbb{V}[P]$$

The framework above assumes we can generate perfect samples of the random variable P. In practice, for complex problems such as SPDEs this is typically impossible. Instead, we must compute numerical approximations which we will denote by P_L, where L represents a level of discretisation. For example, and as will be used later, in a finite difference scheme L could correspond to a mesh with spatial grid spacing $h_L = 2^{-(L+1)}$. A higher L means a finer mesh and a more accurate - but also more computationally expensive - approximation.

This introduces a second source of error. The total error of our estimate is now a combination of the statistical error from the Monte Carlo sampling and the systematic bias from the numerical discretisation. In this context, the MSE of the standard MC estimator which uses N samples of the numerical approximation P_L, which we denote $\hat{P}_{MC,L}$ is therefore the following:

$$\text{MSE}(\hat{P}_{MC,L}) = \underbrace{\frac{1}{N}\mathbb{V}[P_L]}_{\substack{\text{Statistical Error} \\ \text{(Variance)}}} + \underbrace{(\mathbb{E}[P_L] - \mathbb{E}[P])^2}_{\substack{\text{Discretisation Error} \\ \text{(Bias}^2\text{)}}} \tag{2.1}$$

This explicit formula presents the clear dilemma of the MC estimator. To achieve an overall MSE less than ε^2, one must sufficiently reduce the above terms. To make the statistical error small, we require a large number of samples N. To make the discretisation error small, we require a fine discretisation level L such that $P_L \approx P$. A finer level L dramatically increases the computational cost of generating each sample, which we denote by C_L. The total cost of the MC estimator is $N \times C_L$. Thus, we have that to achieve a target MSE, both terms in this product increase.

This is the fundamental challenge that the Multilevel Monte Carlo method is designed to overcome. Instead of estimating the expensive quantity $\mathbb{E}[P_L]$ directly, MLMC rewrites it using a telescoping sum:

$$\mathbb{E}[P_L] = \sum_{\ell=0}^{L} \mathbb{E}[Y_\ell], \quad \text{where} \quad Y_\ell := P_\ell - P_{\ell-1} \quad \text{and} \quad P_{-1} := 0.$$

Each term $\mathbb{E}[Y_\ell]$ is then estimated independently with a standard Monte Carlo estimator.

Definition 3 (Multilevel Monte Carlo Estimator) *Let $Y_\ell = P_\ell - P_{\ell-1}$ be the correction at level ℓ. The **Multilevel Monte Carlo (MLMC) estimator**, $\hat{P}_{\text{MLMC}}$, for the expectation $\mathbb{E}[P_L]$ is:*

$$\hat{P}_{\text{MLMC}} = \sum_{\ell=0}^{L} \hat{Y}_\ell, \quad \textit{where} \quad \hat{Y}_\ell = \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} Y_\ell^{(n)}.$$

By the linearity of expectation, $\hat{P}_{\text{MLMC}}$ is an unbiased estimator for $\mathbb{E}[P_L]$. Since the estimates at each level are independent, its variance is the sum of the individual variances. We define the cost and variance of a single sample at level ℓ as C_ℓ and V_ℓ respectively:

$$
\begin{aligned}
C_\ell &:= \text{Cost}(Y_\ell) \\
V_\ell &:= \mathbb{V}[Y_\ell] = \mathbb{V}[P_\ell - P_{\ell-1}] \\
&= \mathbb{V}[P_\ell] + \mathbb{V}[P_{\ell-1}] - 2\text{Cov}(P_\ell, P_{\ell-1}).
\end{aligned}
$$

The total cost and variance of the MLMC estimator are therefore:

$$
C_{\text{MLMC}} = \sum_{\ell=0}^{L} N_\ell C_\ell, \qquad \mathbb{V}[\hat{P}_{\text{MLMC}}] = \sum_{\ell=0}^{L} \frac{V_\ell}{N_\ell}.
$$

The success of the MLMC method hinges on the behaviour of the level variances, V_ℓ. The key is to ensure that V_ℓ is small, especially for large ℓ where the cost C_ℓ is high. This is achieved by using the same underlying source of randomness to generate the pair $(P_\ell, P_{\ell-1})$. This technique is known as *coupling*. For example, when solving a Stochastic Differential Equation driven by Brownian motion, both the fine (P_ℓ) and coarse $(P_{\ell-1})$ simulations use the same discretised Brownian path ([8], section 5.1 for example).

Because the coarse and fine paths are strongly correlated, functionals P_ℓ and $P_{\ell-1}$ are also strongly correlated. Consequently, the variance of their difference, V_ℓ, is much smaller than the variance of either term individually. As the level ℓ increases, P_ℓ converges to $P_{\ell-1}$, and so we expect that $V_\ell \to 0$.

This property allows for the following trade-off: we can use a small number of samples N_ℓ for the expensive, high-level correction terms (where V_ℓ is small) and compensate by using a large number of samples for the cheap, low-level terms. We can determine the optimal number of samples for each N_ℓ that minimises the total cost in order to achieve a given variance and also ensure we use enough levels such that our discretisation error is also sufficiently small. Both of these ensure that can achieve a target MSE or RMSE at minimal cost.

For a fixed variance ε^2, choosing the optimal $\{N_\ell\}_{\ell=0}^{L}$ that minimises the total cost $C_{\text{MLMC}} = \sum N_\ell C_\ell$ is a classic constrained optimisaton problem solveable with Lagrange mutipliers. It yieds optimal number of samples:

$$
N_\ell = \left\lceil \frac{1}{\varepsilon_{\text{var}}^2} \sqrt{\frac{V_\ell}{C_\ell}} \sum_{k=0}^{L} \sqrt{V_k C_k} \right\rceil. \tag{2.2}
$$

The ceiling function $\lceil \cdot \rceil$ ensures the number of samples is an integer. Equation (2.2) states that we should take more samples when the variance per unit cost of a level is high, and less when it is low.

We also have that the bias $(\mathbb{E}[P_\ell] - \mathbb{E}[P]) \to 0$ as $\ell \to \infty$. To ensure that MSE is less than ε^2, by Definition 2 we can impose that $(\mathbb{E}[P_L - P])^2 < \frac{\varepsilon^2}{2}$ and $\mathbb{V}[\hat{P}_{\text{MLMC}}] < \frac{\varepsilon^2}{2}]$.

This leads to the following theorem [8] which makes precise the cost scaling of the MLMC method:

Theorem 1 (MLMC Complexity Theorem) *Let P denote a random variable, and let P_ℓ denote the corresponding level ℓ numerical approximation.*

If there exists independent estimators Y_ℓ based on N_ℓ Monte Carlo 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 we have

1. ***Weak Error (Bias) Decay:*** $|\mathbb{E}[P_\ell - P]| \leq c_1 2^{-\alpha\ell}$,
2. ***Unbiased Estimators:*** $\mathbb{E}[Y_\ell] = \begin{cases} E[P_0], & l = 0 \\ E[P_l - P_{l-1}], & l > 0, \end{cases}$
3. ***Variance Decay:*** $V_\ell \leq c_2 2^{-\beta\ell}$,
4. ***Cost Growth:*** $C_\ell \leq c_3 2^{\gamma\ell}$,

then there exists a positive constant c_4 such that for any $\varepsilon < e^{-1}$ there are values L and N_L for which the multilevel estimator

$$Y = \sum_{\ell=0}^{L} Y_\ell,$$

has an MSE with bound

$$MSE \equiv \mathbb{E}\left[(Y - \mathbb{E}[P])^2\right] < \varepsilon^2 \tag{2.3}$$

with a computational complexity C with bound

$$\mathbb{E}[C] \leq \begin{cases} c_4\varepsilon^{-2}, & \beta > \gamma, \\ c_4\varepsilon^{-2}(\log\varepsilon)^2, & \beta = \gamma, \\ c_4\varepsilon^{-2-(\gamma-\beta)/\alpha},] & \beta < \gamma \end{cases} \tag{2.4}$$

To appreciate the significance of the MLMC Complexity Theorem, we first establish the cost of the standard MC method in the above context. To achieve an MSE of $\mathcal{O}(\varepsilon^2)$, both the statistical and discretisation errors must be controlled. Controlling the **bias** to $O(\varepsilon)$ requires using a fine grid with a step size $h_L \propto \varepsilon^{1/\alpha}$. Independently, controlling the statistical error to $O(\varepsilon^2)$ requires $N \propto \varepsilon^{-2}$ samples. The total cost

is the product of the number of samples and the cost per sample, where $C_L \propto h_L^{-\gamma}$. Combining these requirements gives the overall complexity:

$$C_{\text{MC}} \propto N \times C_L \propto \varepsilon^{-2} \times (\varepsilon^{1/\alpha})^{-\gamma} = \varepsilon^{-2-\gamma/\alpha}.$$

In contrast, in the MLMC case $\beta > \gamma$, the dominant computational cost is on coarsest level where the cost per sample C_ℓ is $O(1)$ [8]. Requiring $N = O(\varepsilon^{-2})$ samples provides the dominant cost. This is the optimal case.

When $\beta = \gamma$, the cost contribution from each level, $N_\ell C_\ell \propto \sqrt{V_\ell C_\ell}$, is approximately constant across all levels [8]. The total cost is therefore proportional to the number of levels, L, which must increase as $\mathcal{O}(\log \varepsilon)$ to meet the bias requirement. This results in the total complexity of $\mathcal{O}(\varepsilon^{-2}(\log \varepsilon)^2)$, which remains a vast improvement over the standard method.

In the case $\beta < \gamma$, the cost per level grows with ℓ, meaning the total cost is dominated by the work on the finest level, L. Even in this worst case scenario though, we still arrive at a smaller scaling of cost than the MC estimator.

We conclude this scenario highlighting therefore the importance of how β and γ relate to one another in determining the magnitude of cost savings an MLMC implementation can offer over an MC estimator. Determining how MLMC performs relative to MC in SPDE applications will depend heavily on what variance decay we can achieve relative to cost growth in our implementations.

2.2 Stochastic Partial Differential Equations

In their most general sense, an SPDE is a partial differential equation where at least one of the following is random: coefficients, initial boundary conditions, the domain, and the forcing term [13]. In our case, the SPDEs we examine have only a random forcing term. In this section, we will present the two SPDEs we consider, define their relevant terms and discuss their specifics.

2.2.1 The Stochastic Heat Equation

The stochastic heat equation (SHE) is a canonical example of a parabolic SPDE. Formally, it is the standard heat equation perturbed by a stochastic forcing term introducing spatially and temporally uncorrelated fluctuations. For instance, heat diffusing through a metal bar that experiences heat emitting chemical reactions is a typical scenario the SHE describes.

$$\frac{\partial u(t,\mathbf{x})}{\partial t} = \Delta u(t,\mathbf{x}) + \xi(t,\mathbf{x}) \tag{SHE}$$
$$u(0,\mathbf{x}) = u_0(\mathbf{x}) \tag{2.5}$$

$u(t,\mathbf{x})$ is a real valued function dependent on time $t \in [0,T]$ and spatial position $\mathbf{x}$ within domain $D \subset \mathbb{R}^d$. Δ is the Laplacian operator.

The defining component is the stochastic forcing term $\xi(t,\mathbf{x})$ term which denotes space-time white noise. ξ is not a classical function but a generalised stochastic process, or random distribution. It is mmost accurately understand as the distributional derivative of a Brownian sheet, $W(t,\mathbf{x})$, which is a centred Gaussian process indexed by $(t,\mathbf{x})$ with independent increments over disjoint rectangles in space-time. ξ is therefore formally defined by its inner product against a test function ϕ, as $\langle \xi, \phi \rangle$. This random variable is Gaussian and defined by the following two properties:

$$\text{Zero Mean: } \mathbb{E}\left[\langle \xi, \phi \rangle\right] = 0 \tag{2.6}$$
$$\text{Covariance Structure: } \mathbb{E}\left[\langle \xi, \phi \rangle \langle \xi, \psi \rangle\right] = \langle \phi, \psi \rangle_{L^2} = \int_0^T \int_D \phi(t,\mathbf{x})\psi(t,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \tag{2.7}$$

In section SECTION HERE we will derive the finite difference schemes used here for these quantities using equations (2.6).

2.2.2 The Dean-Kawasaki Equation

The Dean-Kawasaki (DK) equation is used to describe the evolution of the density $\rho(\mathbf{x},t)$ of a system of $N >> 1$ weakly interacting particles. For the non-interacting case investigated in this dissertation, the equation is given by:

$$\frac{\partial_\rho(t,\mathbf{x})}{\partial t} = \frac{1}{2}\Delta\rho(t,\mathbf{x}) + N^{-1/2}\nabla \cdot (\sqrt{\rho(t,\mathbf{x})}\xi(t,\mathbf{x}))\ \rho(0,\mathbf{x}) \qquad = \rho_0(\mathbf{x}) \tag{DK}$$

Here, $\rho(t,\mathbf{x})$ is the particle density at time t and position $\mathbf{x}$. The term $\frac{1}{2}\Delta\rho$ describes standard particle diffusion. The stochastic forcing term $N^{-1/2}\nabla \cdot (\sqrt{\rho}\xi)$ models the density fluctuations, where ξ is space-time white noise.

The DK equation presents significant mathematical and numerical challenges. The noise term is multiplicative, its magnitude scaled by the local density through $\sqrt{\rho}$. The equation is highly singular. As shown in [**?**], the only martingale solutions to DK are empirical measures of the underlying particle system. Thus we require

$$\rho(\mathbf{x},t) \equiv \mu_t^N(\mathbf{x}) := N^{-1}\sum_{i=1}^{N}\delta(x - X_i(t)).$$

Cornalba and Fischer [**?**] however do demonstrate that statistical properties of fluctuations around the mean-field limit $r\bar{h}o$ can be simulated. Generally, quantities of interest of the form:

$$Q = \psi\big(N^{1/2}\int(\mu_N^T - \bar{\rho}^T)(\mathbf{x})\phi(\mathbf{x})\mathrm{d}\mathbf{x}\big) \tag{2.8}$$

This can correspond to for example the variances $\mathbb{E}[|N^{1/2}\int($
$mu_T^N - \bar{\rho}^T)(\mathbf{x})\phi(\mathbf{x})\mathrm{d}x|^2]$.

2.3 Literature Review

The MLMC method was formally introduced and popularised by Giles for SDE path simulation in 2008 [7], building on earlier foundational work on multilevel integration by Heinrich starting in 1998 [11]. Early works focussed predominantly on application to SDEs, particularly in computational finance. Other research extended the method's application to a wider range of topics, including various classes of SDEs [2, 15], Lévy processes [10], Numerical Linear Algebra [3], and Reliability theory [4].

MLMC for SPDEs is a more recent and active area of research. A search on Scopus [6] of "Multilevel Monte Carlo Stochastic Partial Differential Equations" at the time of writing returns 88 documents. Changing this to "Multilevel Monte Carlo Parabolic Stochastic Partial Differential Equations" refines this down to 18. Of the work examining MLMC for SPDEs, much of the focus has been on elliptic SPDEs (for example [1, 12, 14]).

This dissertation is concerned with parabolic SPDEs, for which the literature provides a smaller but highly relevant set of foundational papers. This review will focus on three key works that inform the central questions of this research: establishing the theoretical basis for MLMC's efficiency, demonstraing its practical validation, and exporing its application to highly singular equations with different coupling strategies.

Barth, Lang, and Schwab [5] analyse provide a foundational analysis of the convegrence and complexity of the MLMC method for a general class of parabolic SPDEs. Using a Galerkin method in space and a Euler-Maruyama scheme in time, they prove that the MLMC estimator significantly reduces the computational work required to achieve a given accuracy compared to a standard single-level method. Their key result shows that the computational complexity can be reduced from $O(h_L^{-(d+4)})$ for a

standard MC method to nearly $O(h_L^{-(d+2)})$ for MLMC, where d is the spatial dimension and h_L is the finest mesh width. Their work provides theoretical underpinning for the cost savings this dissertation seeks to validate for the stochastic heat and Dean-Kawasaki equations. However, their analysis does not explore the practical performance improvements of different noise coupling strategies.

In [9], Giles and Reisinger provide a practical demonstration of MLMC's performance improvements for a class of parabolic SPDEs arising in financial modelling. The authors develop and analyse a Milstein finite difference scheme, proving it converges with first-order accuracy in time and second-order in space. They demonstrate a concrete reduction in computational complexity from $O(\varepsilon^{-7/2})$ for a standard MC approach to the optimal $O(\varepsilon^{-2})$ for their MLMC implementation, validating this gain through numerical experiments. This work serves as a methodological benchmar, demonstrating how to empirically confirm the theoretical performance gains of the MLMC method.

A very revent and highly relevant contribution is the 2024 paper by Cornalba and Fischer, which develops and analyses an MLMC method specifically for the Dean-Kawasaki equation, one o the two case studies in this dissertation. Their work tackles a highly singular SPDE for which standard MLMC convergence proofs fail. By formulating their analysis in terms of the convergence of probability distributions, they prove that MLMC provides a significant computational improvement over standard MC, provided average particle density is sufficiently large. Crucially, they propose and analyse two distinct noise coupling strategies: a "Fourier coupling" and a "Right-Most Nearest Neighbours (NN) coupling". We build directly on this work in this dissertation by also investigating the Dean-KAwasaki equation and proposing an alternative noise coupling strategy.

NEED TO ADD REGARDING HPC MLMC IMPLEMENTATIONS AND SPDES (believe none done)

Bibliography

[1] Assyr Abdulle, Andrea Barth, and Christoph Schwab. Multilevel monte carlo methods for stochastic elliptic multiscale pdes. *Multiscale Modeling & Simulation*, 11(4):1033–1070, 2013.

[2] Assyr Abdulle and Adrian Blumenthal. Stabilized multilevel monte carlo method for stiff stochastic differential equations. *Journal of Computational Physics*, 251:445–460, 2013.

[3] Juan A Acebrón. A probabilistic linear solver based on a multilevel monte carlo method. *Journal of Scientific Computing*, 82(3):65, 2020.

[4] Louis JM Aslett, Tigran Nagapetyan, and Sebastian J Vollmer. Multilevel monte carlo for reliability theory. *Reliability Engineering & System Safety*, 165:188–196, 2017.

[5] Andrea Barth, Annika Lang, and Christoph Schwab. Multilevel monte carlo method for parabolic stochastic partial differential equations. *BIT Numerical Mathematics*, 53(1):3–27, 2013.

[6] Elsevier. Scopus. https://www.scopus.com, Accessed: August 11, 2025.

[7] Michael B Giles. Multilevel monte carlo path simulation. *Operations research*, 56(3):607–617, 2008.

[8] Michael B Giles. Multilevel monte carlo methods. *Acta numerica*, 24:259–328, 2015.

[9] Michael B Giles and Christoph Reisinger. Stochastic finite differences and multilevel monte carlo for a class of spdes in finance. *SIAM journal on financial mathematics*, 3(1):572–592, 2012.

[10] Michael B Giles and Yuan Xia. Multilevel monte carlo for exponential lévy models. *Finance and Stochastics*, 21(4):995–1026, 2017.

[11] Stefan Heinrich. Monte carlo complexity of global solution of integral equations. *Journal of complexity*, 14(2):151–175, 1998.

[12] Ralf Kornhuber, Christoph Schwab, and Maren-Wanda Wolf. Multilevel monte carlo finite element methods for stochastic elliptic variational inequalities. *SIAM Journal on Numerical Analysis*, 52(3):1243–1268, 2014.

[13] Sergey V Lototsky, Boris L Rozovsky, et al. *Stochastic partial differential equations*, volume 11. Springer, 2017.

[14] Yan Luo and Zhu Wang. A multilevel monte carlo ensemble scheme for random parabolic pdes. *SIAM Journal on Scientific Computing*, 41(1):A622–A642, 2019.

[15] Chang-han Rhee and Peter W Glynn. Unbiased estimation with square root convergence for sde models. *Operations Research*, 63(5):1026–1043, 2015.