ExaQUte

Exascale **Q**uantification of **U**ncertainties for
Technology and Science Simulation

D5.4. REPORT ON MLMC FOR TIME DEPENDENT PROBLEMS

Document information table

Contract number:	800898
Project acronym:	ExaQUte
Project Coordinator:	CIMNE
Document Responsible Partner:	EPFL
Deliverable Type:	Report
Dissemination Level:	PUblic
Related WP & Task:	WP5 & Task 5.4
Status:	Approved

Authoring

Prepared by CIMNE and EPFL				
Authors	Partner	Modified	Version	Comments
Quentin Ayoul-Guilmard	EPFL	All	1.0.0	Development
Sundar Ganesh				
Marc Nuñez	CIMNE			
Riccardo Tosi				
Fabio Nobile	EPFL			Supervision
Riccardo Rossi	CIMNE			

Change Log

Versions	Modified Page/Sections	Comments
1.0	All	Submitted version

Approval

Approved by CIMNE and EPFL				
	Name	Partner	Date	OK
Task leader	Fabio Nobile	EPFL	2020-11-13	✓
Coordinator	Riccardo Rossi	CIMNE	2020-11-13	✓

Executive summary

In this report, we study the use of Multi-Level Monte Carlo (MLMC) methods for time dependent problems. It was found that the usability of MLMC methods depends strongly on whether or not the underlying time dependent problem is chaotic in nature. Numerical experiments are conducted on both simple problems, as well as fluid flow problems of practical interest to the ExaQUte project, to demonstrate this. For the non-chaotic cases, the hypotheses that enable the use of MLMC methods were found to be satisfied. For the chaotic cases, especially the case of high Reynolds' number fluid flow, the hypotheses were not satisfied. However, it was found that correlations between the different levels were high enough to merit the use of multi-fidelity or control-variate approaches. It was also noted that MLMC methods could work for chaotic problems if the time window of analysis were chosen to be small enough. Future studies are proposed to examine this possibility.

Contents

List of Figures

List of Tables

Nomenclature / Acronym list

Acronym	Meaning
QoI	Quantity of Interest
MC	Monte Carlo
MLMC	Multi-Level Monte Carlo
CMLMC	Continuation Multi-Level Monte Carlo
MIMC	Multi-Index Monte Carlo
MFMC	Multi-Fidelity Monte Carlo
AMR	Adaptive Mesh Refinement
CVaR	Conditional Value at Risk
CDF	Cumulative Distribution Function
MSE	Mean Squared Error
CFL	Courant–Friedrichs–Lewy
PDE	Partial Differential Equation
ODE	Ordinary Differential Equation
SDE	Stochastic Differential Equation
DNS	Direct Numerical Simulation
RANS	Reynolds-Averaged Navier–Stokes
LES	Large-Eddy Simulation

1 Introduction

MLMC methods have shown success in the last decade to improve the complexity of expensive UQ problems [5, 8, 15, 16]. However, they have seen limited use for UQ in time-dependent PDEs with random parameters. We mention e.g. [9] where MLMC approaches were demonstrated for the Langevin SDE with a strongly-concave potential. The idea was extended in [6] to non-contractive SDEs. MLMC approaches were also demonstrated successfully for systems of hyperbolic conservation laws in [11] and [12].

The target application area of ExaQUte is in civil engineering involving time-dependent turbulent fluid flow models and the target QoI are typically time averages of some aerodynamic quantity such as force or moment coefficients on a building or structure. In this report, we discuss the possibility of using MLMC methods for such time-dependent problems with the goal of computing time averaged quantities. We also explore the potential for the use of ergodicity to improve the efficiency of computations.

In the application context, we are interested in studying the behavior of a QoI of the following form.

$$Q = \langle \tilde{Q} \rangle_T \coloneqq \frac{1}{T} \int_0^T \tilde{Q}(t) dt, \tag{1}$$

where T can either be a finite time window or can go to infinity, signifying an infinite time average if the limit exists, and $\tilde{Q}(t)$ is the time series of the output QoI of our computational model, containing random parameters (typically inlet wind conditions). Both cases of finite and infinite time horizon have important implications for the use of MLMC methods to make computations more efficient.

MLMC methods aim to improve the complexity of Monte Carlo methods by constructing the QoI Q on a sequence of $L + 1$ meshes of characteristic sizes $h_0 > h_1 > ... > h_L$ used to discretize the underlying system of PDEs, such that $h_l = \delta^{-l} h_0$ where $\delta > 1$. The goal is to estimate $\mathbb{E}[Q_{h_L}]$ on the finest discretization level as an approximation to $\mathbb{E}[Q]$. The linearity of the expectation can be used to rewrite it as

$$\mathbb{E}[Q_{h_L}] = \mathbb{E}[Q_{h_0}] + \sum_{l=1}^{L} \mathbb{E}\left[Q_{h_l} - Q_{h_{l-1}}\right] \tag{2}$$

Each of the expectations is estimated by independent Monte Carlo estimators using a different number of i.i.d. replicas $\{N_l\}_{l=0}^{L}$, giving rise to the following MLMC estimator.

$$\hat{\mu} = \frac{1}{N_0} \sum_{i=1}^{N_0} Q_{h_0}(\omega^{(i,0)}) + \sum_{l=1}^{L} \frac{1}{N_l} \sum_{i=1}^{N_l} \left[Q_{h_l}(\omega^{(i,l)}) - Q_{h_{l-1}}(\omega^{(i,l)})\right] \tag{3}$$

$$= \sum_{l=0}^{L} \frac{1}{N_l} \sum_{i=1}^{N_l} Y_l^{(i,l)}, \tag{4}$$

where we have defined $Y_0^{(i,0)} \coloneqq Q_{h_0}(\omega^{(i,0)})$ and $Y_l^{(i,l)} \coloneqq Q_{h_l}(\omega^{(i,l)}) - Q_{h_{l-1}}(\omega^{(i,l)})$ and $\omega^{(i,l)}$ are all i.i.d. realizations of the underlying noise process inducing randomness in the system. The MSE of the estimator is defined as $\text{MSE}(\hat{\mu}) \coloneqq \mathbb{E}[(\hat{\mu} - \mathbb{E}[Q])^2]$ and is given by

$$\text{MSE}(\hat{\mu}) = (\mathbb{E}[Q_h - Q])^2 + \sum_{l=0}^{L} \frac{\mathbb{V}\text{ar}[Y_l]}{N_l}, \tag{5}$$

where the first and second terms are respectively the bias and statistical error. Under the assumption that there exist positive constants $C_\alpha, \alpha, C_\beta, \beta, C_\gamma, \gamma$ such that the following hold:

$$B_l := |\mathbb{E}\left[Q_{h_l} - Q\right]| \leqslant C_\alpha e^{-l\alpha}, \tag{6a}$$

$$V_l := \mathbb{V}\text{ar}\left[Y_l\right] \leqslant C_\beta e^{-l\beta}, \tag{6b}$$

$$C_l := \text{Cost}(Y_l^{(i)}) \leqslant C_\gamma e^{l\gamma}, \tag{6c}$$

it has been shown in [7] that the cost of the MLMC simulation to obtain a tolerance of ϵ^2 on the MSE, split equally between the bias and statistical error contributions, behaves as

$$\text{Cost}(\hat{\mu}) = \sum_{l=0}^{L} N_l C_l \lesssim \begin{cases} \epsilon^{-2} & \beta > \gamma, \\ \epsilon^{-2}(\log \epsilon)^2, & \beta = \gamma, \\ \epsilon^{-2-(\gamma-\beta)/\alpha}, & \beta < \gamma. \end{cases} \tag{7}$$

when $L = L(\epsilon)$ and $N_l = N_l(\epsilon)$ are optimally chosen. For $\beta > \gamma$, the cost is dominated by Monte Carlo sampling on the coarsest levels. For $\beta = \gamma$, the cost is distributed evenly across levels and for $\beta < \gamma$, the cost is primarily on the finest levels. Even in the worst case, MLMC estimators are an improvement over the standard Monte Carlo complexity of $\epsilon^{-2-\gamma/\alpha}$. During implementation, the number of samples N_l and levels L required to attain a given tolerance with optimal complexity depend strongly on the rates α, β and γ. For some problems, these rates are available from theoretical considerations. However in most applications, they are obtained by fitting on estimates of B_l, V_l and C_l respectively. These estimates can be obtained by running a "screening" phase with a few samples or can be estimated and improved on the fly in a continuous manner. The reader is referred to [8, 14] for detailed descriptions of different MLMC algorithms.

2   MLMC theory for time dependent problems

The main challenge in using MLMC methods for time dependent problems is to retain the correlation in time of the fine and coarse QoI time signals. This strongly depends on the properties of the underlying system of equations. Particularly, this depends on whether the governing equations are chaotic or not and if yes, what the time-scales of decorrelation are. Namely, it is important to retain pathwise correlation of the time signals on both fine and coarse meshes through the time window of analysis.

In this report, we study four problems of interest to demonstrate the effects of the chaos on the levelwise convergence hypotheses in Eqs. (6). Namely, we present two simple oscillators - one non-chaotic and one chaotic, which are proxy models for more complex fluid dynamics problems. We also present two cases of practical relevance to the project - Navier-Stokes equations at a low Reynolds' number and at a higher Reynolds' number that supports turbulence. We demonstrate that for the chaotic problems, the hypotheses of Eqs. (6) no longer hold true due to a lack of pathwise-convergence.

To illustrate the potential pitfalls in using MLMC for time-dependent problems, let us consider a system of SDEs with additive noise given by

$$dX(t) = f(t, X(t))dt + \sigma dW(t), \quad t > 0, \quad X(0) = X^0, \tag{8}$$

where $X(t) \in \mathbb{R}^n$, $\sigma \in \mathbb{R}^{n \times k}$, and $W(t)$ is a $\mathbb{R}^k$-valued standard Wiener process. Relevant to the ExaQUte project and this report is the accurate computation of time averages of some output quantities $\tilde{Q}(X(t))$ with $\tilde{Q} : \mathbb{R}^n \to \mathbb{R}$ a smooth function. Namely, our goal is to compute

$$Q = \langle \tilde{Q}(X(t)) \rangle_T = \frac{1}{T} \int_0^T \tilde{Q}(X(t)) dt. \tag{9}$$

We consider as well a discretized version of Eq. (9) by, for example, the Euler–Maruyama scheme.

$$X_{n+1} = X_n + hf(t_n, X_n) + \sigma \Delta W_n, \quad n = 0, 1, ..., \quad X_0 = X^0, \tag{10}$$

with $\Delta W_n \sim \mathcal{N}(0, h)$ i.i.d. (where h is the step size) and a piecewise linear reconstruction of the solution

$$X_h(t) = \left(\frac{t_{n+1} - t}{h} \right) X_n + \left(\frac{t - t_n}{h} \right) X_{n+1}, \quad t \in (t_n, t_{n+1}], \quad n = 0, 1, ... \tag{11}$$

This leads to the following approximation of the time average.

$$Q_h = \langle \tilde{Q}(X_h(t)) \rangle_T = \frac{1}{T} \int_0^T \tilde{Q}(X_h(t)) dt. \tag{12}$$

Under reasonable assumptions on the Hölder continuity of f and $\tilde{Q}$, it is well known [10, Theorem 14.1.5] that

$$\left| \mathbb{E}\left[\tilde{Q}(X_h(t)) \right] - \mathbb{E}\left[\tilde{Q}(X(t)) \right] \right| \leqslant c_1(t)h, \quad \forall t > 0, \tag{13}$$

$$\mathbb{E}\left[\left(\tilde{Q}(X_h(t)) - \tilde{Q}(X(t)) \right)^2 \right]^{1/2} \leqslant c_2(t)h, \quad \forall t > 0. \tag{14}$$

The left hand side of Eq. (13) and Eq. (14) are called the weak error and strong (or pathwise) error, respectively. For general SDEs with Lipschitz coefficients, the strong error of the Euler-Maruyama scheme scales as $\mathcal{O}\left(h^{1/2} \right)$ whereas in the current setting of additive noise, we have rather that it scales as $\mathcal{O}(h)$. Also, assuming that $f(t, \cdot)$ and $\tilde{Q}(\cdot)$ are at least Lipschitz continuous, the constant $c_2(t)$ typically has the form $c_2(t) = \bar{c}_2 e^{Lt}$. For chaotic systems, the constant L is typically positive and large, meaning that the error estimate is meaningful only for a time horizon T of the order $\mathcal{O}\left(L^{-1} \right)$. On the other hand, under reasonable assumptions[1] on f, one has that $\mathbb{E}\left[\tilde{Q}(X(t))^2 \right]^{1/2} \leqslant C$ for all $t > 0$ so that Eq. (14) can be replaced by

$$\mathbb{E}\left[\left(\tilde{Q}(X_h(t)) - \tilde{Q}(X(t)) \right)^2 \right]^{1/2} \leqslant \min \left\{ \bar{c}_2 e^{Lt} h, 2C \right\}, \quad \forall t > 0. \tag{15}$$

Concerning the weak error, if both the SDE and its discretized form are ergodic, the constant $c_1(t)$ is uniformly bounded in time so that Eq. (13) can be replaced by

$$\left| \mathbb{E}\left[\tilde{Q}(X_h(t)) \right] - \mathbb{E}\left[\tilde{Q}(X(t)) \right] \right| \leqslant \bar{c}_1 h, \quad \forall t > 0. \tag{16}$$

---

[1]E.g. $X^\top f(t, X) \leqslant -k_1 |X|^2 + k_2$ for $|X| \geqslant R$ for suitable non-negative constants k_1, k_2 and R.

The estimates in Eqs. (15) and (16) can be used to estimate the decay of the bias and variance estimates B_l and V_l defined in Eqs. (6) on the time averaged QoI Q_{h_l}. We have that

$$B_l = \mathbb{E}\left[Q_{h_l} - Q\right] = \mathbb{E}\left[\langle \tilde{Q}(X_{h_l}(t))\rangle_T - \langle \tilde{Q}(X(t))\rangle\right] \leqslant \bar{c}_1 h_l = \hat{c}_1 \delta^{-l}, \tag{17}$$

$$V_l = \mathbb{V}\mathrm{ar}\left[Q_{h_l} - Q_{h_{l-1}}\right] \leqslant \mathbb{E}\left[\left(Q_{h_l} - Q_{h_{l-1}}\right)^2\right] \tag{18}$$

$$\leqslant 2\,\mathbb{E}\left[(Q_{h_l} - Q)^2\right] + 2\,\mathbb{E}\left[\left(Q_{h_{l-1}} - Q\right)^2\right] \tag{19}$$

$$\leqslant 4\min\left\{4C^2, \frac{\bar{c}_2^2 e^{2LT}}{2L}h_{l-1}^2\right\} \tag{20}$$

$$\leqslant \min\{\hat{c}_2, \hat{c}_3 e^{2LT}\delta^{-2l}\}, \tag{21}$$

for suitable time independent constants $\hat{c}_1, \hat{c}_2, \hat{c}_3$. We see from these estimates that the bias term always features an exponential decay with respect to the level l, whereas to observe a decay of the variance in the chaotic case, we have to either consider very large l or $T = \mathcal{O}(L^{-1})$ so that the second term in the minimum of Eq. (21) is smaller than the first one. In other terms, variance decay should be expected only for short time intervals in which the two approximated outputs $\tilde{Q}(X_{h_l}(t))$ and $\tilde{Q}(X_{h_{l-1}}(t))$ remain correlated. For long time horizons, the two time series decorrelate completely and the variance term V_l, while remaining bounded, does not feature any decay with respect to l.

3 Results for Oscillator Problems

3.1 Van der Pol Oscillator

The Van der Pol Oscillator is an oscillator whose trajectory $x(t)$ is governed by the second-order differential equation

$$\frac{d^2x}{dt^2} - \mu(1 - x^2)\frac{dx}{dt} + x = 0. \tag{22}$$

The oscillator has the favorable property that it has a limit cycle to which it converges independent of the initial coordinates in the phase space. It is also a good proxy model for vortex shedding fluid flows. To assess the effectiveness of MLMC methods in this case, we consider a stochastic version of Eq. (22) where the right hand side is replaced by white noise (derivative of a Wiener process). We are interested in studying the convergence properties of the time average of the trajectory $x(t)$. We first analyze the convergence properties of $\langle x\rangle_T$ in the unforced case. We rewrite Eq. (22) as a set of first order ODEs by introducing the variable $y = dx/dt$.

$$\frac{dx}{dt} = y, \quad t \in (0, T], \tag{23}$$

$$\frac{dy}{dt} = \mu(1 - x^2)y - x, \quad t \in (0, T], \tag{24}$$

$$x(0) = x_0, \quad y(0) = y_0. \tag{25}$$

We discretize this system with the explicit Euler time stepping scheme on a time grid given by $0 = t_0 < t_1 < \cdots < t_N = T$ where $t_n = nh$ and $h = T/N$. This leads to the

following time stepping scheme.

$$\begin{bmatrix} y_{n+1} \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} y_n \\ x_n \end{bmatrix} + h \begin{bmatrix} \mu(1 - x_n^2)y_n - x_n \\ y_n \end{bmatrix}. \tag{26}$$

The time average in Eq. (1) is approximated as

$$\langle x \rangle_T \approx \langle x \rangle_{h,T} := \frac{1}{N} \sum_{n=1}^{N} x_n \tag{27}$$

For the current study, we select $\mu = 1$, $x_0 = 1$ and $y_0 = 1$. We select a time window of $T = 100$ which corresponds to approximately 15 periods of the limit cycle. We carry out the simulation for 5 different values of N, namely 16000, 32000, 64000, 128000 and 256000 points. A plot of the phase space trajectory of the solution for the finest mesh is shown in Fig. 1 The finest mesh is selected as a reference value $\langle x \rangle_{ref,T}$. In Fig. 2, we plot the values $|\langle x \rangle_{h,T} - \langle x \rangle_{ref,T}|$ for different values of h. We observe a convergence rate of approximately 1 based on a least squares fit, which is consistent with the order of the Euler scheme.
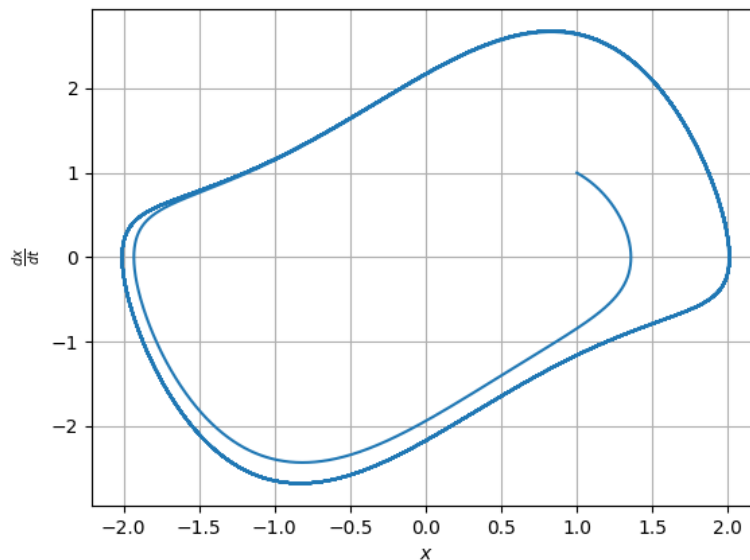

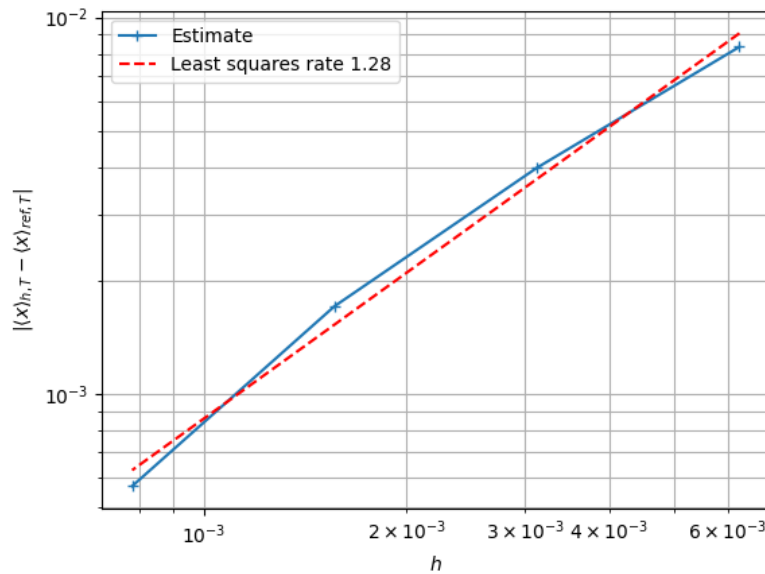Figure 1: Phase space trajectory of Van der Pol oscillator

Figure 2: Mesh convergence in the deterministic case for the Van der Pol oscillator

For the stochastic study, we first look at the pathwise convergence of the forced Van der Pol oscillator given by

$$\frac{d^2x}{dt^2} - \mu(1-x^2)\frac{dx}{dt} + x = \tau\dot{W}(t), \tag{28}$$

where $W(t)$ is a standard Wiener process and τ is the strength parameter for the forcing term. Similar to earlier, we rewrite this equation as a system of first order ODEs.

$$\frac{dx}{dt} = y, \quad t \in (0, T] \tag{29}$$

$$\frac{dy}{dt} = \mu(1-x^2)y - x + \tau\dot{W}(t), \quad t \in (0, T] \tag{30}$$

$$x(0) = x_0, \quad y(0) = y_0. \tag{31}$$

We discretize the system using the Euler-Maruyama scheme, which reads as follows.

$$\begin{bmatrix} y_{n+1} \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} y_n \\ x_n \end{bmatrix} + h \begin{bmatrix} \mu(1-x_n^2)y_n - x_n + \frac{\tau}{\sqrt{h}}\xi_n \\ y_n \end{bmatrix}, \tag{32}$$

where ξ_n are i.i.d. realizations of a standard normal random variable. All parameters remain identical as to the deterministic case. The strength parameter is chosen as $\tau = 1.0$ and the time horizon T is chosen to be 100. The solution $x(t)$ is plotted versus t for the same realization of the white noise solved on both finest and coarsest meshes in Fig. 3. Pathwise correlation can clearly be observed in the plot.
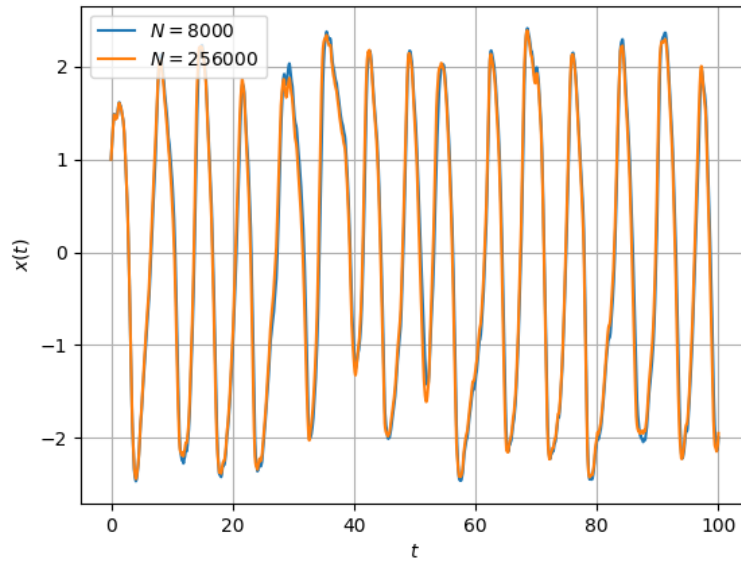
Figure 3: One realization of the stochastic Van der Pol oscillator solution on the finest and coarsest meshes

We simulate the system for 10 independent realizations of the Brownian path. We plot the values $|\langle x \rangle_{h,T}^{(i)} - \langle x \rangle_{ref,T}^{(i)}|$ vs. h for $i = \{1, ..., 10\}$, where for each realization of the Brownian path, the finest mesh is taken to be the reference solution. Each color corresponds to a different underlying Brownian path realization. The resultant plot is shown in Fig. 4. We observe convergence rates of between 1.0 to 1.3 based on least squares fits, which is again consistent with the predicted strong convergence rate of the Euler-Maruyama scheme.
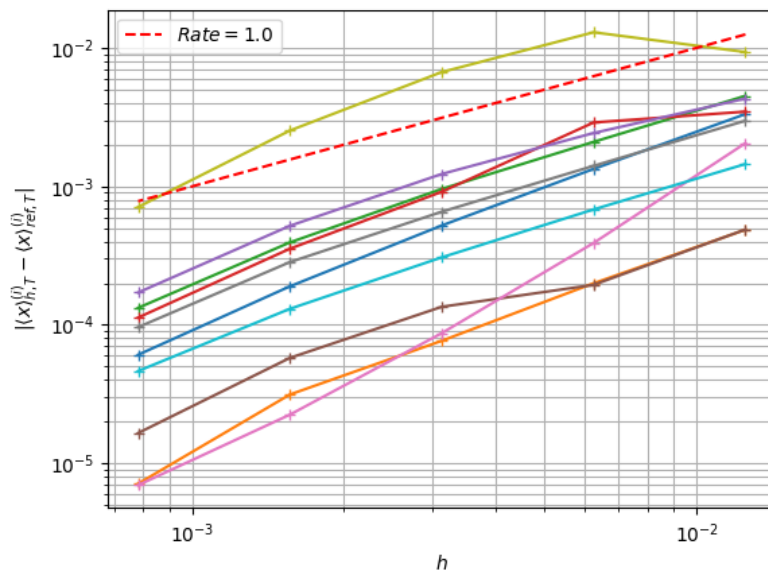
Figure 4: Pathwise mesh convergence in the stochastic case for the Van der Pol oscillator

We wish to study the applicability of MLMC algorithms to this problem. For MLMC to produce the optimal complexity for a given tolerance, the underlying problem should satisfy the rate hypotheses of Eqs. (6). We keep the system parameters the same as earlier. Indexing the meshes in decreasing step size as $l = \{0, ..., L\}$, we denote the quantity of interest on mesh l as $\langle x \rangle_{l,T}$ and the corresponding step size as h_l. We study the convergence of the quantities

$$B_l := |\, \mathbb{E}\left[\langle x \rangle_{l,T} - \langle x \rangle_{ref,T}\right]\,|, \tag{33a}$$

$$V_l := \mathbb{V}\mathrm{ar}\left[\langle x \rangle_{l,T} - \langle x \rangle_{l-1,T}\right], \tag{33b}$$

with respect to h_l. For each l, we estimate the expectation and variance using their corresponding sample average and sample variance estimators using 100 independent Brownian path realizations. For each Brownian path, the problem is solved on both the fine and coarse levels. The variation of the bias and variance with levels is shown in Fig. 5. We observe rates of approximately 1 and 2 for the bias and variance respectively.
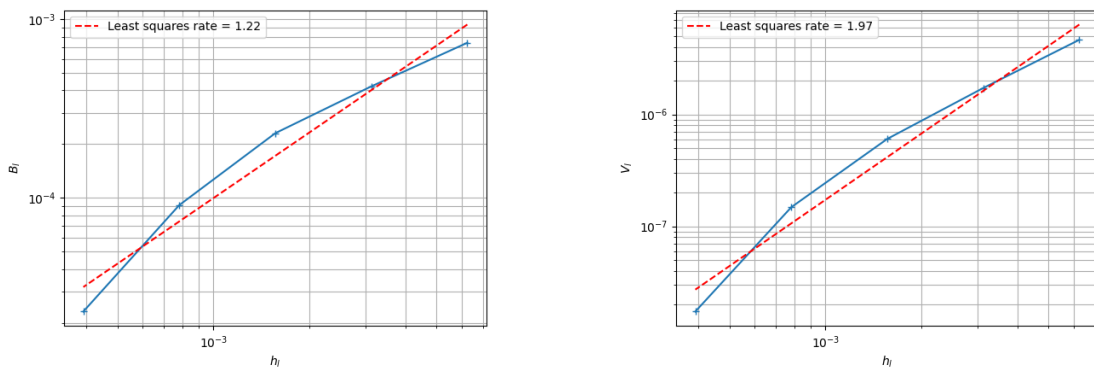
Figure 5: Bias (left) and variance (right) convergence for the stochastic Van der Pol oscillator

We observe that the stochastic Van der Pol oscillator possesses favorable properties in terms of retaining pathwise correlations. In addition, we have also shown that the Van der Pol oscillator satisfies the MLMC hypotheses Eqs. (6) and hence can obtain the demonstrated complexity behavior with an optimally selected hierarchy. To demonstrate this optimal complexity behavior, we estimate the expectation $\mathbb{E}\left[\langle x \rangle_T\right]$ using an optimally tuned MLMC estimator $\hat{\mu}$. The input uncertainty is the Brownian forcing term. We prescribe a tolerance ϵ on the total error of the MLMC estimator defined as

$$\mathrm{TE}(\hat{\mu}) := |\, \mathbb{E}\left[\langle x \rangle_{L,T} - \langle x \rangle_{L-1,T}\right]\,| + C_\alpha \sqrt{\sum_{l=0}^{L} \frac{\mathbb{V}\mathrm{ar}\left[\langle x \rangle_{l,T} - \langle x \rangle_{l-1,T}\right]}{N_l}}, \tag{34}$$

where C_α corresponds to the inverse of the CDF of the standard normal distribution at significance $1 - \alpha/2$.

We use the CMLMC algorithm [5, 16] to tune the hierarchy optimally for a given tolerance on the total error. The cost of computing the optimally tuned hierarchy is then measured and plotted against the corresponding tolerance. For each tolerance tested, the entire MLMC simulation is repeated 15 times and the corresponding simulation time is

noted for the optimally tuned hierarchy. The results are shown in Fig. 6. It can be seen that the cost grows as ϵ^{-2}, thus demonstrating optimal complexity behavior for CMLMC. In addition, the estimated cost for a Monte Carlo simulation to reach the same tolerance is also shown.
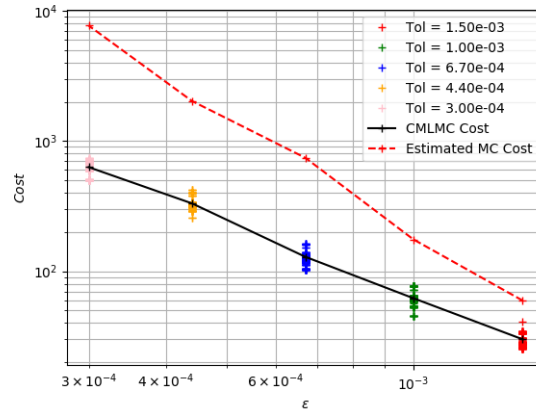
Figure 6: Complexity behavior for CMLMC algorithm on Van der Pol oscillator

3.2 Lorenz Oscillator

The Lorenz oscillator is a three dimensional chaotic oscillator governed by the following system of ODEs.

$$\frac{dx}{dt} = \sigma(y - x), \tag{35a}$$

$$\frac{dy}{dt} = x(\rho - z) - y, \tag{35b}$$

$$\frac{dz}{dt} = xy - \beta z. \tag{35c}$$

For the purposes of this study, we select the parameter values to be $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$. The Lorenz oscillator has the property that it is chaotic. This means that two trajectories with the initial conditions differing by an arbitrarily small perturbation will eventually diverge. This poses a challenge for MLMC since pathwise convergence is important for the hypotheses to be satisfied.

As before, we study two variations of the oscillator. The first is the pathwise convergence of the deterministic oscillator in Eqs. (35). The second is a stochastic oscillator where the right hand sides of all three of Eqs. (35) are forced with independent white noises as follows,

$$\frac{dx}{dt} = \sigma(y - x) + \tau \dot{W}_1(t), \tag{36a}$$

$$\frac{dy}{dt} = x(\rho - z) - y + \tau \dot{W}_2(t), \tag{36b}$$

$$\frac{dz}{dt} = xy - \beta z + \tau \dot{W}_3(t), \tag{36c}$$

for $t \in (0, T]$, where T denotes the time horizon, W_1, W_2 and W_3 are independent Wiener processes and τ is the strength parameter whose value is chosen to be 1.0 for the purposes of this study.

For the first study, we discretize Eqs. (35) using the explicit Euler scheme on a uniform grid with $0 = t_0 < t_1 < \cdots < t_N = T$ with $t_n = nh$ and $h = T/N$. The discretized system reads

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \\ z_{n+1} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} + h \begin{bmatrix} \sigma(y_n - x_n) \\ x_n(\rho - z_n) - y_n \\ x_n y_n - \beta z_n \end{bmatrix}. \tag{37}$$

We analyze the system for $T = 400$ and for different values of N, namely - $\{4, 8, 16, 32, 64\} \times 10^4$. Considering the finest mesh as a reference, we plot the value of the differences $|\langle x \rangle_{h,T} - \langle x \rangle_{ref,T}|$ for different values of h. The resultant plot is shown in Fig. 7. We note that we do not observe convergence of the quantity of interest in this case.
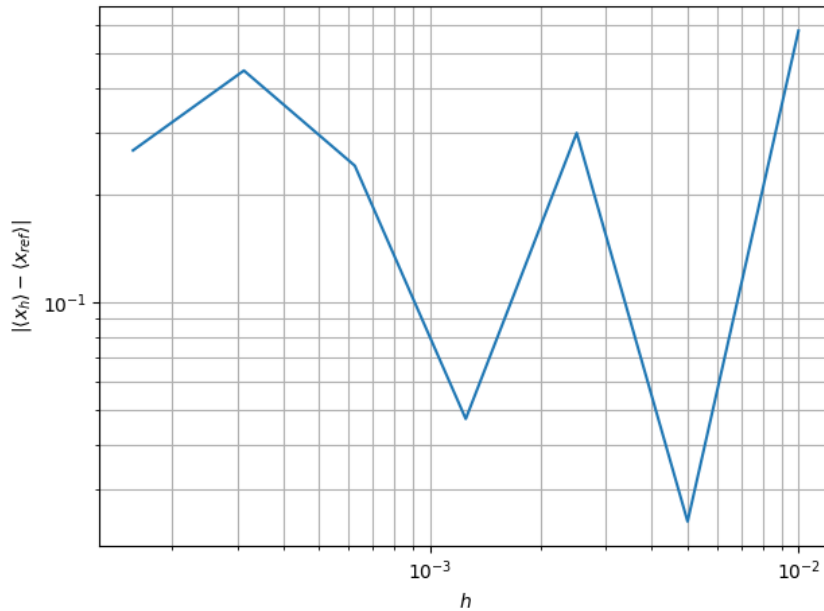
Figure 7: Convergence of time averages for the deterministic Lorenz oscillator

For the second study, we discretize Eqs. (36) using the Euler-Maruyama scheme. The discretized system reads

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \\ z_{n+1} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} + h \begin{bmatrix} \sigma(y_n - x_n) + \frac{\tau}{\sqrt{h}}\xi_{1,n} \\ x_n(\rho - z_n) - y_n + \frac{\tau}{\sqrt{h}}\xi_{2,n} \\ x_n y_n - \beta z_n + \frac{\tau}{\sqrt{h}}\xi_{3,n} \end{bmatrix}, \tag{38}$$

where $\xi_{1,n}, \xi_{2,n}$ and $\xi_{3,n}$ are independent standard normally distributed random variables. Fig 8 shows one realization of the solution of the forced Lorenz oscillator computed on the finest and coarsest meshes with the same underlying white-noise realizations. We observe that the solutions very quickly decorrelate.
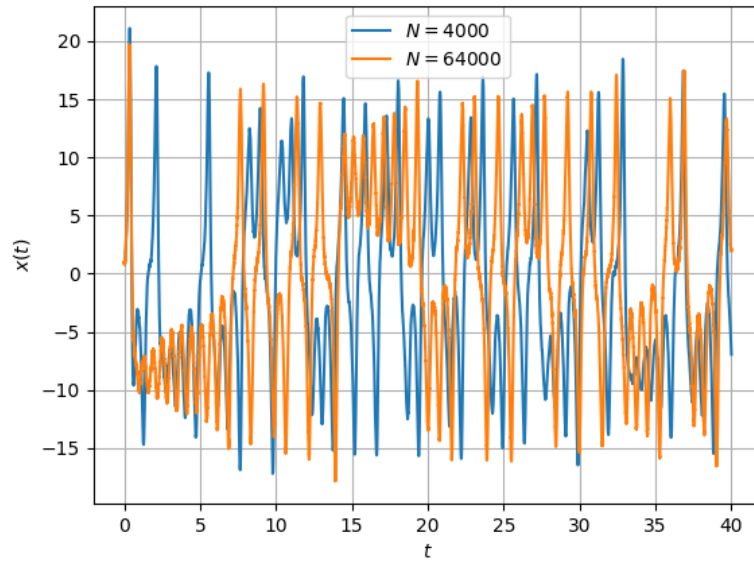
Figure 8: A realization of the stochastic Lorenz oscillator solution computed on the finest and coarsest meshes

Similar to before, we simulate the system for 10 independent realizations of the Brownian paths. We plot the values $|\langle x \rangle^{(i)}_{h,T} - \langle x \rangle^{(i)}_{ref,T}|$ vs. h for $i = \{1, ..., 10\}$, where for each realization of the Brownian path, the finest mesh is taken to be the reference solution. The resultant plot is shown in Fig. 9 where different colors indicate the different realizations.
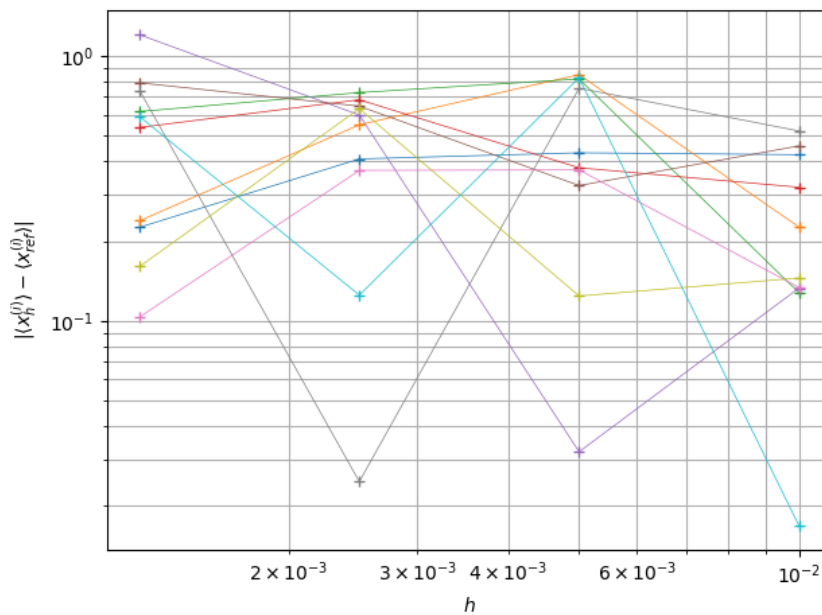
Figure 9: Convergence of time averages for the stochastic Lorenz oscillator

Lastly, we conduct a screening MLMC similar to Section 3.1 with 10^4 samples per pair of levels. We study the decay of the bias and variance defined in Eqs. (33) with levels l.

It is expected that the bias decays for a large enough time window T but that variance decay cannot be guaranteed. The resultant behavior is reported in Fig. 10 and it can be observed that although the bias decays with rate better than 1 in h, the variance does not decay at all.
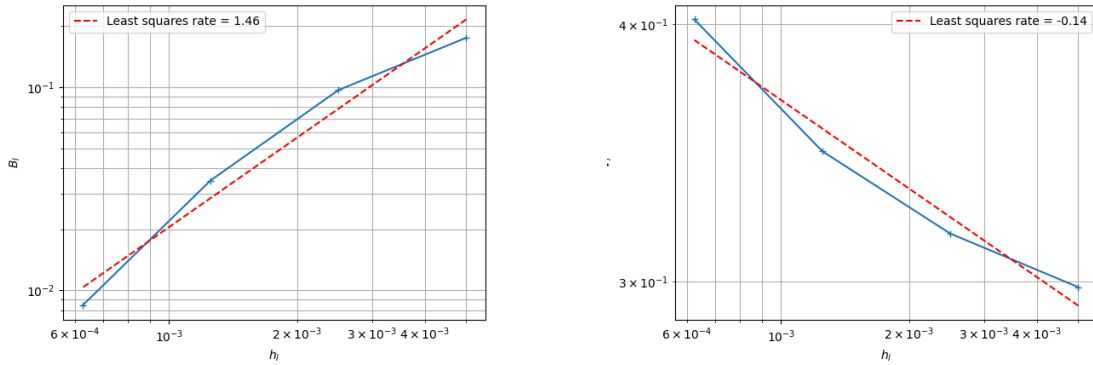
Figure 10: Bias (left) and Variance (right) convergence for the stochastic Lorenz oscillator

4 Results for Fluid Flow Problems

4.1 Low Reynolds' Number Fluid Flow

In the ExaQUte project, we aim to simulate the fully turbulent flow over a three dimensional building. As an intermediate case, we consider a reduced problem in two dimensions - the flow over a 5×1 rectangle. A schematic of the problem is shown in Fig. 11 and Fig. 12.
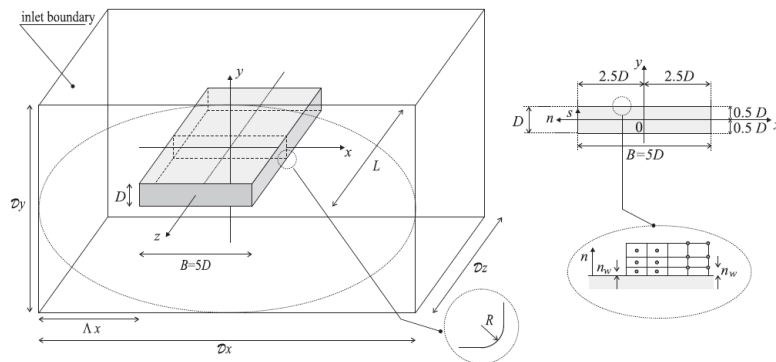
Figure 11: Problem description [4], $D = 1$

The incompressible Navier-Stokes equations for the velocity field u and pressure field p are used to model the fluid flow:

$$\frac{\partial u}{\partial t} - \nu \Delta u + (u \cdot \nabla)u + \nabla p = 0, \tag{39}$$

$$\nabla \cdot u = 0; \tag{40}$$
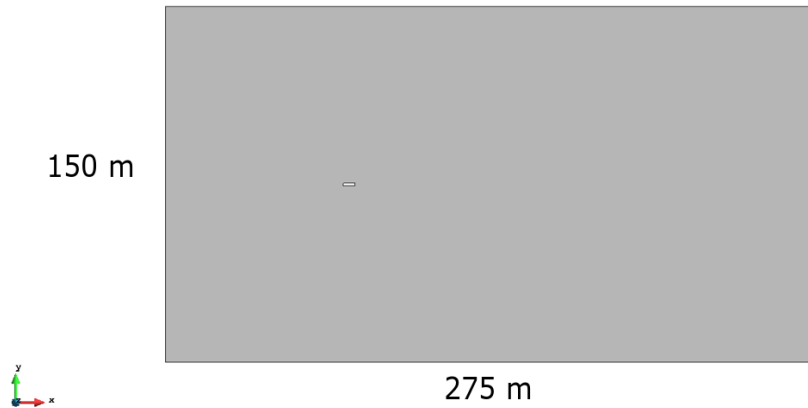
Figure 12: Rectangle problem dimensions. Inner rectangle 5×1.

and the QoI considered is the time-averaged drag force. The details of the case are described extensively in [1]. It was shown in ibid. that pathwise convergence could be obtained at a relatively low Reynolds' number of 100 if the correct adaptive mesh strategy was used. The details of the resultant meshes can be found in Appendix A. The problem was solved on a background mesh for an inlet velocity of 2.0 m/s and the resultant velocity and pressure fields were used to refine and create a hierarchy of meshes. A description of the adaptive strategy can also be found in ibid. We refer to ibid. for the results on the convergence study.

This test case however converges to a steady solution. We are currently investigating the more interesting case of $Re = 150$ which exhibits vortex shedding and limit cycle oscillations on both lift and drag forces. We will consider a normally distributed random inlet velocity. We plan to explore bias and variance convergence for this problem and also to combine it with continuation type MLMC algorithms [5] to iteratively improve the meshes based on information from all random samples.

4.2   High Reynolds' Number Fluid Flow

For the study of higher Reynolds' number, we present here the results from [1]. The same numerical example as in Section 4.1 is simulated but with the viscosity adjusted to yield a higher Reynolds' number. The time-averaged velocity field is used to compute metrics for mesh adaptivity. The parameters of the resultant meshes can be seen in Table 1. The simulations are done for $Re \approx 1.3 \times 10^5$. The inlet velocity is made stochastic. A

Interpolation Error	h_{min}	Nodes [$\times 1000$]	CFL	h
10^1	0.035	1.1	80	0.7
10^0	0.012	2	80	0.24
10^{-1}	0.0033	5	80	0.066
10^{-2}	0.0011	15	80	0.022
10^{-3}	0.00037	92	80	0.0075

Table 1: Mesh parameters for high Reynolds' number study

total of 50 realizations of the inlet velocity $v_{inlet} \sim \mathcal{N}(2.0, 0.02)$ are simulated on all of

the meshes, which are designed with the 'mean' case of $v_{inlet} = 2.0$. It is of interest to study the convergence of $|\langle F_{D,l}^{(i)} \rangle - \langle F_{D,l-1}^{(i)} \rangle|, i \in \{1, 2, ..., 50\}$ for each of the samples, where $\langle F_D, l \rangle$ denotes the time averaged value of the drag force F_D over the the time window $[140, 300]$. We measure the convergence of this quantity versus the interpolation error prescribed to the adaptive remesher in order to obtain each level of mesh.

Fig. 13 shows the variation of $|\langle F_{D,l}^{(i)} \rangle - \langle F_{D,l-1}^{(i)} \rangle|$ with the interpolation error for each of the different realizations, as well as $\mathbb{E}[\langle F_{D,l} \rangle - \langle F_{D,l-1} \rangle]$ estimated using a sample average over the 50 realizations. It can be seen from the plot that geometric decay of this quantity in the mesh is not observed. The variation of the variance of the differences
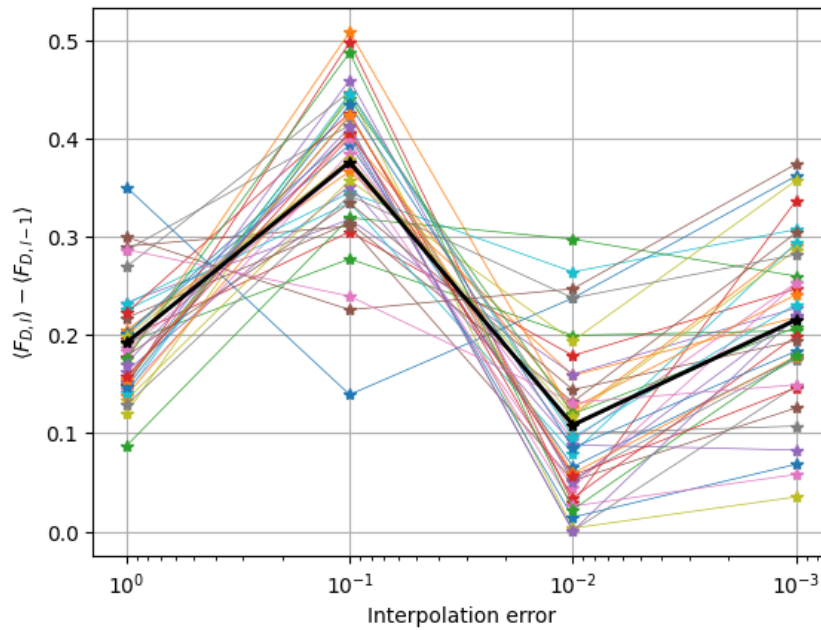
Figure 13: Bias decay plot for the flow problem with high Reynolds number. Same color for same realization at different levels. Black line denotes sample average. Levels defined by Table 1.

$\mathbb{V}ar[\langle F_{D_l} \rangle - \langle F_{D_{l-1}} \rangle]$ is also plotted in Fig. 14. As can be seen from the figure, this quantity does not decay with the mesh parameters either.
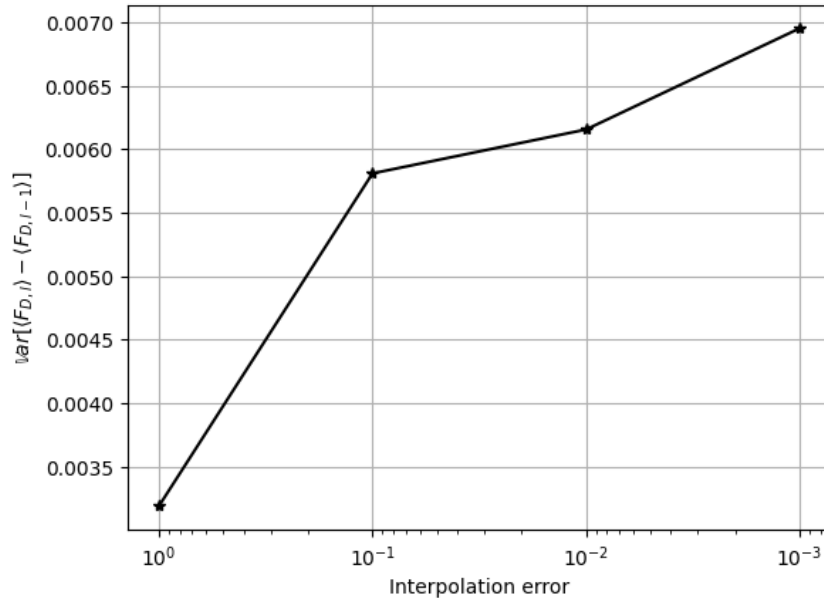
Figure 14: Variance decay plot for the flow problem with high Reynolds number. Levels defined by Table 1

On further investigation, it was found that bias and variance decay could be difficult to obtain at high Reynolds' numbers due to the chaotic nature of the flow. This makes it very difficult to retain pathwise correlation of both fine and coarse samples. As a result, the MLMC hypotheses of Eqs. (6) are likely to not be fulfilled. In the following sections, we outline some promising ideas and algorithms for increasing the efficiency of computations

5   Recommendations and Future Scope

5.1   Multi-Fidelity Monte Carlo Methods

From the data available from the studies in Section 4.2, it is possible to compute cross-level correlations. Although pathwise-convergence was not observed, the correlations observed suggested the use of MFMC methods to provide significant speed-ups. The correlation data is presented in Table 3. We present here the basic theory of the methods of control variates and MFMC.

Let X be a random variable with bounded second moment. We recall the definition of a MC estimator of $\mathbb{E}[X]$ with m samples:

$$\mu_m(X) := \frac{1}{m} \sum_{i=1}^{m} X^{(i)}, \tag{41}$$

where $X^{(i)})_{i=1}^m$ are i.i.d. replicas of X. This is an unbiased estimator of $\mathbb{E}[X]$, therefore its mean squared error is

$$\mathbb{E}\left[(\mu_m(X) - \mathbb{E}[X])^2\right] = \mathbb{V}\mathrm{ar}[\mu_m(X)] = \frac{\mathbb{V}\mathrm{ar}[X]}{m}. \tag{42}$$

A straightforward way to improve the accuracy of the Monte Carlo estimator is to increase m. In our applications, this is often infeasible. An alternative method is to "reduce" $\mathbb{V}\mathrm{ar}\,[X]$ by substituting to X a random variable $\tilde{X}$ of smaller variance and equal mean value.

We introduce here the idea of variance reduction by control variates. Let Z be another random variable with bounded second moment. Z serves as a control variate by defining the controlled variable $\tilde{X} = X - \alpha(Z - \mathbb{E}\,[Z])$. We have, purposely, $\mathbb{E}\left[\tilde{X}\right] = \mathbb{E}\,[X]$. The variance $\mathbb{V}\mathrm{ar}\left[\tilde{X}\right] = \mathbb{V}\mathrm{ar}\,[X] - 2\alpha\,\mathbb{C}\mathrm{ovar}\,(X, Z) + \alpha^2\,\mathbb{V}\mathrm{ar}\,[Z]$ is minimal for

$$\alpha = \frac{\mathbb{C}\mathrm{ovar}\,(X, Z)}{\mathbb{V}\mathrm{ar}\,[Z]}, \tag{43}$$

with minimum

$$\min_{\alpha \in \mathbb{R}} \mathbb{V}\mathrm{ar}\left[\tilde{X}\right] = \left(1 - \frac{\mathbb{C}\mathrm{ovar}\,(X, Z)^2}{\mathbb{V}\mathrm{ar}\,[X]\,\mathbb{V}\mathrm{ar}\,[Z]}\right) \mathbb{V}\mathrm{ar}\,[X]. \tag{44}$$

We observe from equation (44) that, whatever Z is, a variance reduction can always be achieved with a suitable choice of the parameter α. A good control variate should be highly correlated to X: the higher the correlation between X and Z, the greater the variance reduction. However, $\mathbb{E}\,[Z]$ must be precisely known lest the control variate introduces a bias. Assuming that it is not known a priori, we propose to approximate $\mathbb{E}\left[\tilde{X}\right] = \mathbb{E}\,[X - \alpha Z] + \alpha\,\mathbb{E}\,[Z]$ with the unbiased estimator

$$\mu_{m_1, m_2}(Z, X) := \frac{1}{m_2} \sum_{i=1}^{m_2} (X^{(i)} - \alpha Z^{(i)}) + \frac{\alpha}{m_1} \sum_{i=1}^{m_1} Z^{(i)}, \tag{45}$$

where $\forall i \in \mathbb{N}$, $X^{(i)} \sim X$ and $Z^{(i)} \sim Z$; and with $m_2 \in \mathbb{N}_*$ and $m_1 > m_2$. Note the *subsampling*: the first m_2 samples of Z appear twice in (45). Therefore, the variance of this estimator is[2]

$$\mathbb{V}\mathrm{ar}\,[\mu_{m_1, m_2}(Z, X)] = \frac{\mathbb{V}\mathrm{ar}\,[X]}{m_2} + \frac{m_1 - m_2}{m_1 m_2}\left(\alpha^2\,\mathbb{V}\mathrm{ar}\,[Z] - 2\alpha\,\mathbb{C}\mathrm{ovar}\,(X, Z)\right). \tag{46}$$

The cost of the estimator is $\textnormal{\euro}\,(\mu_{m_1, m_2}(Z, X)) := m_2\,\textnormal{\euro}\,(X) + m_1\,\textnormal{\euro}\,(Z)$. Additionally, the benefit of an optimal α comes at the cost of estimating $\mathbb{V}\mathrm{ar}\,[Z]$ and $\mathbb{C}\mathrm{ovar}\,(X, Z)$.

Let us generalize estimator (45) to any number of control variates. We have a set of random variables $\mathbf{X} := (X_i)_{i=1}^{M+1}$, alternatively called "models". X_{M+1} is the "high-fidelity" model whose expectation we want to estimate (formerly X), and the others are M "low-fidelity" models used as control variates. Our MFMC estimator is defined as

$$\mu_{\mathbf{m}}(\mathbf{X}) := \frac{1}{m_{M+1}} \sum_{i=1}^{m_{M+1}} \left(X_{M+1}^{(i)} - \sum_{j=1}^{M} \alpha_j X_j^{(i)}\right) + \sum_{j=1}^{M} \frac{\alpha_j}{m_j} \sum_{i=1}^{m_j} X_j^{(i)}, \tag{47}$$

with $\mathbf{m} \in \mathbb{N}^{M+1}$ samples chosen such that $\forall i \in [\![1, M]\!]$, $m_i > m_{M+1} > 0$. There, again, we use subsampling so that for every pair $\{j, k\} \subset [\![1, M]\!]$ and for any $i \in \mathbb{N}$, $X_j^{(i)}$ and $X_k^{(i)}$ use the same realization of the underlying noise process and are therefore correlated.

---

[2]This is a particular case of the expression obtained in appendix B.

To express its variance, let us introduce the following notations:

$$\mathbf{r} := \left(\frac{m_j}{m_{M+1}} \right)_{j \in [\![1,M]\!]}, \tag{48}$$

$$R := \operatorname{diag} \mathbf{r}, \tag{49}$$

$$C := \left(\frac{\mathbb{C}\mathrm{ovar}\,(X_i, X_j)}{\mathbb{V}\mathrm{ar}\,[X_{M+1}]} \right)_{i,j \in [\![1,M]\!]}, \tag{50}$$

$$\mathbf{c} := \left(\frac{\mathbb{C}\mathrm{ovar}\,(X_{M+1}, X_j)}{\mathbb{V}\mathrm{ar}\,[X_{M+1}]} \right)_{j \in [\![1,M]\!]}, \tag{51}$$

$$\tilde{C} := (C_{i,j}(\min\{r_i, r_j\} - 1))_{i,j \in [\![1,M]\!]}, \tag{52}$$

and

$$A := (R^{-1} - I)C(R^{-1} - I) + R^{-1}\tilde{C}R^{-1}, \tag{53}$$

with the identity operator

$$I := (\delta_{i,j})_{i,j \in [\![1,M]\!]}. \tag{54}$$

Then, the values of the control coefficients which minimize the variance of the MFMC estimator are

$$\alpha^{\star} := A^{-1}(1 - R^{-1})\mathbf{c} \tag{55}$$

for which

$$\mathbb{V}\mathrm{ar}\,[\mu_{\mathbf{m}}(\mathbf{X})] = \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \left(1 - \mathbf{c}^{\top}(R^{-1} - 1)A^{-1}(R^{-1} - 1)\mathbf{c} \right). \tag{56}$$

The detailed proof of (55) and (56) is presented in appendix B, p. 33. For a fixed budget B, we define the optimal allocation of samples $\check{\mathbf{m}}$ as

$$\check{\mathbf{m}} := \left\lfloor \arg\min \left\{ \mathbb{V}\mathrm{ar}\,[\mu_{\mathbf{m}}(\mathbf{X})] : \right. \right.$$
$$\left. \left. \mathbf{m} \in \,]0, +\infty[^{M+1}; \; \forall i \in [\![1, M]\!], \, m_i > m_{M+1}; \; \mathbb{\in}(\mu_{\mathbf{m}}(\mathbf{X})) \leqslant B \right\} \right\rfloor \tag{57}$$

where $\mathbb{\in}(\mu_{\mathbf{m}}(\mathbf{X})) = \sum_{i=1}^{M+1} m_i \mathbb{\in}(X_i)$.

Note that no assumption has been made regarding the comparative accuracy of the low-fidelity models. In the case where the models can be ordered in a hierarchy satisfying some assumptions, [13] has proposed a different class of estimators with explicit formulæ for optimal control coefficients $\{\alpha_i : i \in [\![1, M]\!]\}$ and sample allocation $\{m_i : i \in [\![1, M]\!]\}$.

We wish to assess the efficacy of the previous MFMC approach on the results presented in Section 4.2. This dataset contains time series of the drag force F_D generated on several meshes for the same set of 50 realizations of the inlet velocity. We consider the random variable

$$Q_i(\omega) := \frac{1}{160} \int_{140}^{300} F_D^{(i)}(t, \omega)\, \mathrm{d}t \tag{58}$$

where $i \in \{1, 2, 3, 4\}$ refers to one of our selected models, which correspond to discretization on different meshes. Table 2 describes these models, of which Q_4 is the high-fidelity one.

Model	Nodes	Mesh size	Timestep (s)	Average cost (s)
1	1100	0.035	0.7	106
2	2000	0.012	0.24	423
3	5000	0.0033	0.066	2821
4	15 000	0.0011	0.022	26 123

Table 2: Models

We will assess the variance reduction that the MFMC estimator presented above can provide compared to a plain MC estimator with the same budget. From formula (56), the data we need are the cost of each model and the covariance of each pair of models. Both are computed from the 50 samples that we have for each model; the average cost measured is displayed in table 2 while the estimated covariance matrix is:

$$\mathbb{C}\text{ovar}\left(\mathbf{Q}\right) = \begin{pmatrix} 0.707 & 0.726 & 0.512 & 0.833 \\ 0.726 & 2.17 & 0.559 & 1.098 \\ 0.512 & 0.559 & 1.391 & 0.64 \\ 0.833 & 1.098 & 0.64 & 2.312 \end{pmatrix} \times 10^{-3}. \tag{59}$$

For better insights, we look at the correlation coefficient matrix presented in Table 3.

Correlations	1	2	3	4
1	1			
2	0.586	1		
3	0.516	0.322	1	
4	0.652	0.490	0.357	1

Table 3: Correlation data for high Reynolds' number fluid flow

To assess the performance of the MFMC estimator, we compute the optimal sample allocation for a fixed budget $B := 50 \times \mathbb{C}\left(Q_4\right)$. Then we calculate the variance reduction compared to the MC estimator with the same budget:

$$\frac{\mathbb{V}\text{ar}\left[\mu_{50}(Q_4)\right]}{\mathbb{V}\text{ar}\left[\mu_{\breve{\mathbf{m}}}(\mathbf{Q})\right]}. \tag{60}$$

Finally, for insight into cost reduction, we provide the factor f by which the number of samples of the MC estimator would have to be multiplied to achieve the same variance reduction, i.e. $\mathbb{V}\text{ar}\left[\mu_{f \times 50}(Q_4)\right] = \mathbb{V}\text{ar}\left[\mu_{\breve{\mathbf{m}}}(\mathbf{Q})\right]$. All these results are summarized in table 4.

Note that some of the low-fidelity models are actually detrimental to the MFMC estimator, due to their poor correlation. The performance of the MFMC estimator was improved by removing them, as is visible in lines "MFMC–3" and "MFMC–2" of Table 4. Lastly, Algorithm 1 describes a tentative MFMC algorithm to calibrate the estimator in Eq. (47).

Estimator	Samples per model				Variance reduction	Equivalent MC sample factor
	1	2	3	4		
Monte Carlo				50	1	1
MFMC–4	356	44	44	43	0.751	1.332
MFMC–3	531	50		46	0.681	1.468
MFMC–2	319			48	0.657	1.522

Table 4: Variance reduction at fixed budget

Algorithm 1 MFMC algorithm to compute optimal number of samples for a given budget

Require: Budget B; screening sample number m'; models $\{X_i : i \in [\![1, M+1]\!]\}$.
 1: Conduct m' simulations each of the $M+1$ correlated models.
 2: Estimate the covariances and variances in Eqs. (50) and (51) using their respective sample estimators.
 3: Compute the optimal control coefficients in Eq. (55)
 4: Compute the optimal sample numbers in Eq. (57) for budget B.
 5: Simulate the additional required samples.

5.2 Burn-in Time Reduction

During the computation of time averages such as the one presented in Eq. (1), it is common practice to leave out a first time interval since the trajectory in this interval is heavily influenced by the initial conditions. This corresponds to computing the QoI as

$$Q_b = \langle \tilde{Q} \rangle_{T, T_{bt}} = \frac{1}{T - T_{bt}} \int_{T_{bt}}^{T} \tilde{Q}(t) dt. \tag{61}$$

The window T_{bt} that is discarded is known as the burn-in time. It can be shown that for geometrically ergodic systems, the bias error of the estimator Q_b is reduced by a factor exponentially decaying in T_{bt} with respect to the bias of Q (without burn-in). However, adequately reducing the bias can imply that a large burn-in time needs to be discarded. In addition, solutions can be computed in parallel in space but not in time, making it important to reduce overall the number of time steps that are computed. Lastly, for ergodic systems, multiple samples can be computed with relatively shorter time windows instead of one sample with a long time window. However, for each sample, the time average of the output QoI has to be computed only after the burn-in time is complete in order to collect statistics with respect to the invariant measure. To make computations more efficient, it is hence important to reduce the burn-in time.

To this end, studies were conducted for the rectangle problem for higher Reynolds' numbers to assess the burn-in time characteristics and propose strategies for its reduction. Instead of the inlet perturbations described in Sec. 4.2, multiple realizations are created by randomly perturbing the initial conditions on the velocity field. A pre-computed velocity field, which is the time-averaged velocity field of a reference solution, is loaded as the initial condition. Perturbations are computed in a manner that the resultant velocity field is divergence-free and observes no-slip conditions on all surfaces.

The QoI in consideration is the time average of the drag force $Q = \langle F_d \rangle_{T,T_{bt}}$. The sample average estimator for the expected value of the time average then reads

$$\mathbb{E}\left[Q\right] \approx \hat{\mu} = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{T - T_{bt}} \int_{T_{bt}}^{T} F_d^{(i)}(t) dt, \tag{62}$$

where $F_d^{(i)}$ is the time series of the output QoI for the i$^{\text{th}}$ realization of the initial velocity. The statistical error SE of this estimator can be estimated as

$$\text{SE} := \mathbb{V}\text{ar}\left[\hat{\mu}\right] \approx \frac{\mathbb{V}\left[F_d\right]}{M(T - T_{bt})^2} \int_{T_{bt}}^{T} \int_{T_{bt}}^{T} \rho(t - s) dt ds, \tag{63}$$

where $\rho(t-s)$ denotes the autocorrelation function of the time series and $\mathbb{V}[F_d]$ is defined as

$$\mathbb{V}\left[F_d\right] := \langle (F_d - \langle F_d \rangle_{T,T_{bt}})^2 \rangle_{T,T_{bt}}. \tag{64}$$

In the earlier studies the burn-in time was set approximately equal to the washout time $T_{bt} = 140\,\text{s}$ to ensure converged computations. We try to analyze how the statistical results of the time-averaged drag force would change with different burn-in times. We plot the expected value $\mathbb{E}\left[\langle F_d \rangle_{T,T_{bt}}\right]$ as a function of T_{bt}, together with its confidence interval. Such intervals are computed as $C_\alpha \sqrt{\text{SE}}$, similar to Eq. (34), with a confidence of 99%. As we can see in Fig. 15a, from a statistical point of view, the statistical error is relatively insensitive to T_{bt} for $T_{bt} \geqslant 20$. However, we decide to apply a physical constraint, which is the time the flow need to go from the inlet to the body. Such time, for an average speed of $2\,\text{m/s}$, is $40\,\text{s}$. As a result, we decided to take $T_{bt} = 40\,\text{s}$ when the system is initialized as described above.

Another way to estimate T_{bt} is following the same approach as in [3], in which the authors choose a burn-in time which minimizes the estimated variance of the sample average estimator of the time average for a given signal. For this reason, we average at each time step over all realizations for a different numbers of realizations and apply the procedure to the resultant time signal. We report the results of this analysis in Fig. 15b for different numbers of realizations. As can be seen from the plot, we reach a minimum value after relatively few seconds. From these results, we assume that it is safe to reduce T_{bt} to $40\,\text{s}$ without affecting the statistical properties of the signal.
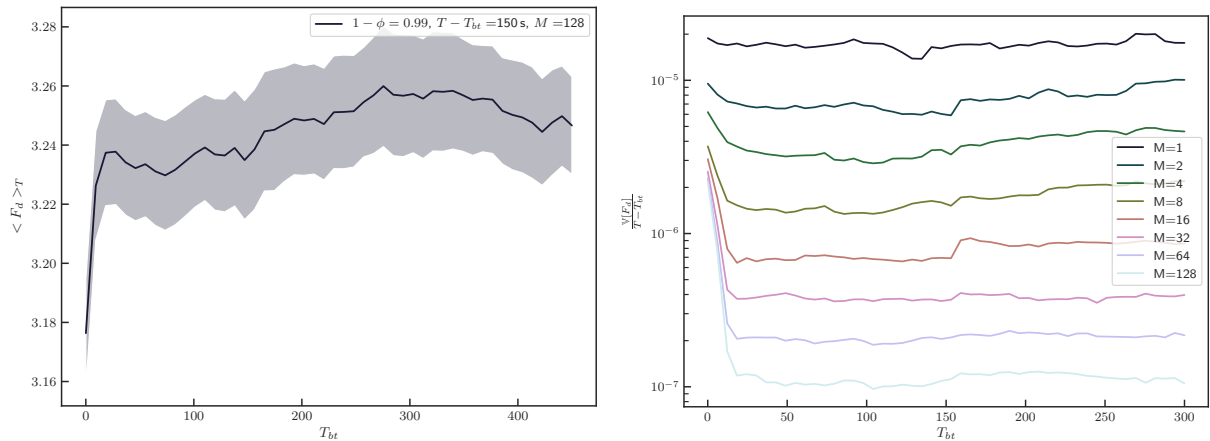
(a) Expected value and associated statistical error for a confidence of 99%.

(b) Ratio between drag force variance and effective time window, for different realizations.

Figure 15: Summary of burn in time reduction experiments.

Another approach we consider to reduce the time to solution is to exploit a larger CFL number during the burn-in time. To do so, we must first verify that a larger CFL number does not change the statistical results. We should also verify that the chosen larger time step, if used during the whole time window, would give different results.

We report in Table 5 the expected value and the associated statistical error for a confidence of 99%. We can see that in the case of time step $h = 0.02$ used in the averaging window $[T_{bt}, T]$, we obtain consistent statistical results independent of the tested values of h_{bt}. On the other hand, for a different h we obtain a different statistical result. Therefore, a larger time step can safely be employed to reduce the time to solution of the burn-in phase.

$\mathbb{E}\left[\langle F_d \rangle_T\right]$	$C_\phi\sqrt{\text{SE}}$	T	T_{bt}	h	h_{bt}
3.237284	0.009183	900	140	0.02	0.02
3.232869	0.008286	900	140	0.02	0.05
3.312648	0.008199	900	140	0.05	0.05

Table 5: Expected value and associated statistical error with 99% confidence for different time steps during both burn-in and effective phases.

5.3 MLMC for Short Time Windows

It was seen from earlier numerical experiments that for chaotic systems, some correlation could still be expected for time windows that were shorter than the decorrelation time. Since pathwise correlation can be expected within such short time windows, it is possible that the simulation can benefit from MLMC approaches. The challenge is then to reduce or eliminate the burn-in time such that the random process is very close to its stationary distribution from the initial time. In this way, the bias error can be eliminated and all samples within the relatively short time window can be used for statistical estimation.

To initialize the problem very close to its stationary distribution, one can run a few simulations on coarser meshes until convergence. An oracle can then be trained on information from this converged solution. This oracle can then be used to propose candidates

for initializing the finer simulations to a state that is close to the stationary distribution. This can be achieved using any of a number of different approaches - reduced basis methods, neural networks, etc. We aim to explore such a method for the case of higher Reynolds' numbers since we can still expect some correlation on the temporal and spatial scales introduced by inlet forcing.

A relatively simple set of studies is planned to first assess the correlations introduced in lift and drag time signals through random inlet forcing with certain spatial and temporal correlations. As long as the scales of the forcing are larger than those of turbulent fluctuations in the fluid, we expect to retain some correlation between finer and coarser meshes since we still expect both simulations to exhibit similar macroscopic structures.

6 Software release

In addition to this report, deliverable 5.4 comprises a new version of the library for hierarchical Monte Carlo methods developed for the ExaQUte project. This version 2.0.0 of the XMC library has been publicly released online [2]. Below is a summary of the major changes with respect to the previous release, i.e. version 1.0.0.

Added

- Updating moment estimators in small batches.
- Asynchronous algorithms for single- and multi-level Monte Carlo methods.
- Estimation of "combined" expectation (time and events) for random processes.
- Support for multi-valued (e.g. vectors) random variables.
- Estimation of moments of order 3 and 4, with a posteriori estimation of statistical error.
- Examples both built-in and using Kratos as external solver.
- Test cases.
- New features of the Kratos interface:
  - different meshing strategies (support distributed environments);
  - write to file in distributed environments.

Changed

- Documentation re-created anew, updated and extended; available in `HTML` format.
- The choice of framework for distributed computing is contained in a single module definition.

Fixed

- Excessive memory usage in case of high number of estimators.

7 Conclusions

In this report, we studied the use of MLMC methods for forward UQ in time dependent problems. It was found that the hypotheses Eqs. (6) were strongly affected by whether or not the underlying problem was chaotic in nature. Studies on two simple oscillators and two fluid problems of practical interest to the ExaQUte project were conducted to illustrate this.

It was shown that the MLMC method could be successfully used for the Van der Pol oscillator, which is a non-chaotic oscillator with a limit cycle. Pathwise convergence was observed and as a result, bias and variance convergence could be obtained at the theoretically predicted rates. A successful CMLMC simulation was also demonstrated that calibrates the MLMC estimator to obtain the optimal complexity behavior.

For the Lorenz oscillator, theoretical considerations predicted that bias convergence could be obtained for long time windows but that variance convergence would be difficult to obtain given that pathwise convergence cannot be maintained for chaotic systems with large Lyapunov exponents. Both predictions were numerically confirmed with the bias decaying at a rate predicted by the theoretical considerations.

For the fluid flow problems, we presented results from [1], wherein mesh convergence was observed for a Reynolds' number of 100 and steady solutions with an appropriately chosen time step size and adaptive meshing strategy. We aim to study the case with a Reynolds' number of 150 since we expect to see vortex shedding at this frequency. It is expected that given the limit cycle behavior of the system, we will be able to benefit from the use of MLMC methods to accelerate the simulations.

For the higher Reynolds' number case, pathwise convergence could not be obtained given the chaotic nature of the fluid flow. However, the resultant correlations were found to be large enough to benefit from the use of MFMC and control variate methods. Future studies are planned to further assess these algorithms and their potential benefits. In addition, strategies were proposed to sample an initial state close to the limit distribution so as to dramatically reduce the burn-in time that is discarded for computing time averages. Lastly, it is planned to explore the use of shorter time windows of analysis in the problem. Theoretical considerations predict that the use of short time windows (in comparison to the correlation time scales) can lead to variance decay as well.

References

[1] Q. Ayoul-Guilmard, S. Ganesh, M. Nuñez, R. Tosi, F. Nobile, and R. Rossi. D5.3. report on theoretical work to allow the use of mlmc with adaptive mesh refinement. deliverable, ExaQUte, 2020.

[2] Q. Ayoul-Guilmard, S. Ganesh, F. Nobile, R. Rossi, R. Tosi, R. Amela, and R. M. Badia. XMC: library for hierarchical monte carlo methods. DOI:`10.5281/zenodo.4265429`, 2020-11-10. version 2.0.0.

[3] P. Beyhaghi, S. Alimohammadi, and T. Bewley. A multiscale, asymptotically unbiased approach to uncertainty quantification in the numerical approximation of infinite time-averaged statistics. *arXiv e-prints*, arXiv:1802.01056, feb 2018.

[4] L. Bruno, M. V. Salvetti, and F. Ricciardelli. Benchmark on the aerodynamics of a rectangular 5: 1 cylinder: an overview after the first four years of activity. *Journal of Wind Engineering and Industrial Aerodynamics*, 126:87–106, 2014.

[5] N. Collier, A.-L. Haji-Ali, F. Nobile, E. Von Schwerin, and R. Tempone. A continuation multilevel monte carlo algorithm. *BIT Numerical Mathematics*, 55(2):399–432, 2015.

[6] W. Fang and M. B. Giles. Multilevel monte carlo method for ergodic sdes without contractivity. *Journal of Mathematical Analysis and Applications*, 476(1):149–176, 2019.

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

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

[9] M. B. Giles, M. B. Majka, L. Szpruch, S. Vollmer, and K. Zygalakis. Multi level monte carlo methods for a class of ergodic stochastic differential equations. *arXiv preprint*, arXiv:1605.01384, 2016.

[10] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 2013.

[11] S. Mishra and C. Schwab. Sparse tensor multi-level monte carlo finite volume methods for hyperbolic conservation laws with random initial data. *Mathematics of Computation*, 81(280):1979–2018, 2012.

[12] S. Mishra, C. Schwab, and J. Šukys. Multi-level monte carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions. *Journal of Computational Physics*, 231(8):3365–3388, 2012.

[13] B. Peherstorfer, K. Willcox, and M. Gunzburger. Optimal model management for multifidelity monte carlo estimation. *SIAM Journal on Scientific Computing*, 38(5): A3163–A3194, 2016.

[14] M. Pisaroni. Multi level monte carlo methods for uncertainty quantification and robust design optimization in aerodynamics. Technical report, EPFL, 2017.

[15] M. Pisaroni, S. Krumscheid, and F. Nobile. Quantifying uncertain system outputs via the multilevel Monte Carlo method — Part I: central moment estimation. Technical Report 23.2017, MATHICSE, 2017.

[16] M. Pisaroni, F. Nobile, and P. Leyland. A continuation multi level monte carlo (C-MLMC) method for uncertainty quantification in compressible inviscid aerodynamics. *Computer Methods in Applied Mechanics and Engineering*, 326:20–50, 2017.

A Mesh Parameters

Table 6: List of meshes used in adaptive refinement study for the flow over a rectangle at $Re = 100$. Error scaled in each approach to accomplish comparable number of nodes and minimal mesh size.

Interpolation error	Minimal mesh size	Number of nodes
225	0,001206	2655
112,5	0,00093	3947
56,25	0,000619	6516
28,125	0,000434	11624
14,0625	0,000249	21838
7,03125	0,000183	42042
3,515625	0,000128	82410
1,757812	0,000101	162998
0,878906	0,000069	322375

B Variance of MFMC estimator

Let us recall the definition of the MFMC estimator introduced in (47),

$$\mu_{\mathbf{m}}(\mathbf{X}) := \frac{1}{m_{M+1}} \sum_{i=1}^{m_{M+1}} \left(X_{M+1}^{(i)} - \sum_{j=1}^{M} \alpha_j X_j^{(i)} \right) + \sum_{j=1}^{M} \frac{\alpha_j}{m_j} \sum_{i=1}^{m_j} X_j^{(i)}, \tag{65}$$

with $\mathbf{m} \in \mathbb{N}^{M+1}$ samples such that $\forall i \in [\![1, M]\!]$, $m_i > m_{M+1} > 0$. We also recall the following notations:

$$\mathbf{r} := \left(\frac{m_j}{m_{M+1}} \right)_{j \in [\![1,M]\!]} \tag{66}$$

$$R := \operatorname{diag} \mathbf{r} \tag{67}$$

$$C := \left(\frac{\mathbb{Covar}(X_i, X_j)}{\mathbb{Var}[X_{M+1}]} \right)_{i,j \in [\![1,M]\!]} \tag{68}$$

$$\mathbf{c} := \left(\frac{\mathbb{Covar}(X_{M+1}, X_j)}{\mathbb{Var}[X_{M+1}]} \right)_{j \in [\![1,M]\!]} \tag{69}$$

$$\tilde{C} := (C_{i,j}(\min\{r_i, r_j\} - 1))_{i,j \in [\![1,M]\!]}, \tag{70}$$

and the identity operator

$$I := (\delta_{i,j})_{i,j \in [\![1,M]\!]}. \tag{71}$$

Let us rearrange the right side of (65) into two independent groups:

$$\mu_{\mathbf{m}}(\mathbf{X}) = \underbrace{\sum_{i=1}^{m_{M+1}} \left(\frac{X_{M+1}^{(i)}}{m_{M+1}} + \sum_{j=1}^{M} \left(\frac{1}{m_j} - \frac{1}{m_{M+1}} \right) \alpha_j X_j^{(i)} \right)}_{(72a)} + \underbrace{\sum_{j=1}^{M} \frac{\alpha_j}{m_j} \sum_{i=1+m_{M+1}}^{m_j} X_j^{(i)}}_{(72b)}. \tag{72}$$

From the independence of samples, we have that

$$\mathbb{Var}[\mu_{\mathbf{m}}(\mathbf{X})] = \mathbb{Var}[(72a)] + \mathbb{Var}[(72b)]. \tag{73}$$

We consider (72a) first.

$$\mathbb{Var}[(72a)] = \frac{\mathbb{Var}[X_{M+1}]}{m_{M+1}} + \mathbb{Var}\left[\sum_{i=1}^{m_{M+1}} \left(\sum_{j=1}^{M} \left(\frac{1}{m_j} - \frac{1}{m_{M+1}} \right) \alpha_j X_j^{(i)} \right) \right]$$
$$- 2 m_{M+1} \mathbb{Covar}\left(\frac{X_{M+1}}{m_{M+1}}, \sum_{j=1}^{M} \left(\frac{1}{m_{M+1}} - \frac{1}{m_j} \right) \alpha_j X_j \right). \tag{74}$$

Let us develop separately the last two terms of the right side of (74):

$$\mathbb{Var}\left[\sum_{i=1}^{m_{M+1}} \left(\sum_{j=1}^{M} \left(\frac{1}{m_j} - \frac{1}{m_{M+1}} \right) \alpha_j X_j^{(i)} \right) \right] \tag{75}$$

$$= m_{M+1} \sum_{j,k \in [\![1,M]\!]} \alpha_j \alpha_k \left(\frac{1}{m_j} - \frac{1}{m_{M+1}} \right) \left(\frac{1}{m_k} - \frac{1}{m_{M+1}} \right) \mathbb{Covar}(X_j, X_k) \tag{76}$$

$$= \frac{\mathbb{Var}[X_{M+1}]}{m_{M+1}} \sum_{j,k \in [\![1,M]\!]} \alpha_j \alpha_k (1 - r_j^{-1})(1 - r_k^{-1}) C_{j,k} \tag{77}$$

$$= \frac{\mathbb{Var}[X_{M+1}]}{m_{M+1}} \alpha^\top (I - R^{-1}) C (I - R^{-1}) \alpha, \tag{78}$$

and

$$- 2m_{M+1} \, \mathbb{C}\mathrm{ovar} \left(\frac{X_{M+1}}{m_{M+1}}, \sum_{j=1}^{M} \left(\frac{1}{m_{M+1}} - \frac{1}{m_j} \right) \alpha_j X_j \right) \tag{79}$$

$$= -2 \sum_{j=1}^{M} \left(\frac{1}{m_{M+1}} - \frac{1}{m_j} \right) \alpha_j \, \mathbb{C}\mathrm{ovar} \left(X_{M+1}, X_j \right) \tag{80}$$

$$= -2 \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \sum_{j=1}^{M} \alpha_j (1 - r_j^{-1}) c_j \tag{81}$$

$$= -2 \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \alpha^\top (I - R^{-1}) \mathbf{c}. \tag{82}$$

Consequently,

$$\mathbb{V}\mathrm{ar}\,[(72\mathrm{a})] = \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \Big(1 - 2\alpha^\top (I - R^{-1}) \mathbf{c}$$
$$+ \alpha^\top (I - R^{-1}) C (I - R^{-1}) \alpha \Big). \tag{83}$$

Secondly, we consider (72b).

$$\mathbb{V}\mathrm{ar}\,[(72\mathrm{b})] = \mathbb{C}\mathrm{ovar} \left(\sum_{j=1}^{M} \frac{\alpha_j}{m_j} \sum_{i=1+m_{M+1}}^{m_j} X_j^{(i)}, \sum_{k=1}^{M} \frac{\alpha_k}{m_k} \sum_{i=1+m_{M+1}}^{m_k} X_k^{(i)} \right) \tag{84}$$

$$= \sum_{j,k \in [\![1,M]\!]} \frac{\alpha_j \alpha_k}{m_j m_k} (\min\{m_j, m_k\} - m_{M+1}) \, \mathbb{C}\mathrm{ovar} \left(X_j, X_k \right) \tag{85}$$

$$= \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \sum_{j,k \in [\![1,M]\!]} \frac{\alpha_j \alpha_k}{r_j r_k} (\min\{r_j, r_k\} - 1) C_{j,k} \tag{86}$$

$$= \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \alpha^\top R^{-1} \tilde{C} R^{-1} \alpha \tag{87}$$

We put together (83) and (87) and get

$$\mathbb{V}\mathrm{ar}\,[\mu_{\mathbf{m}}(\mathbf{X})] = \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \Big(1 - 2\alpha^\top (I - R^{-1}) \mathbf{c} \tag{88}$$
$$+ \alpha^\top \underbrace{\left((I - R^{-1}) C (I - R^{-1}) + R^{-1} \tilde{C} R^{-1} \right)}_{A} \alpha \Big).$$

$$=: \nu(\alpha) \tag{89}$$

Furthermore,

$$\arg\min \left\{ \nu(\alpha) : \alpha \in \mathbb{R}^M \right\} = A^{-1} (I - R^{-1}) \mathbf{c} =: \alpha^\star \tag{90}$$

and

$$\nu(\alpha^\star) = \frac{\mathbb{V}\mathrm{ar}\,[X_{M+1}]}{m_{M+1}} \left(1 - \mathbf{c}^\top (I - R^{-1}) A^{-1} (I - R^{-1}) \mathbf{c} \right). \tag{91}$$