ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Institute of Mathematics

School of Basic Sciences

MATHICSE Technical Report

Nr. 23.2017

October 2017

Quantifying uncertain system outputs via the multilevel Monte Carlo method – Part I: Central moment estimation

Michele Pisaroni, Sebastian Krumscheid, Fabio Nobile

http://mathicse.epfl.ch

Address: EPFL - SB – INSTITUTE of MATHEMATICS – Mathicse
(Bâtiment MA) Station 8 - CH-1015 - Lausanne - Switzerland

Quantifying uncertain system outputs via the multilevel Monte Carlo method — Part I: Central moment estimation

M. Pisaroni[a,*], S. Krumscheid[a], F. Nobile[a]

[a] *Calcul Scientifique et Quantification de l'Incertitude (CSQI), Institute of Mathematics, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland*

Abstract

In this work we introduce and analyze a novel multilevel Monte Carlo (MLMC) estimator for the accurate approximation of central moments of system outputs affected by uncertainties. Central moments play a central role in many disciplines to characterize a random system output's distribution and are of primary importance in many prediction, optimization, and decision making processes under uncertainties. We detail how to effectively tune the MLMC algorithm for central moments of any order and present a complete practical algorithm that is implemented in our accompanying Python library `cmlmc-py`[1]. In fact, we validate the methodology on selected reference problems and apply it to an aerodynamic relevant test case, namely the transonic RAE 2822 airfoil affected by operating and geometric uncertainties.

Keywords: central moments, multilevel Monte Carlo, uncertainty quantification, h-statistic

1. Introduction

Probabilistic models are widely used across many disciplines, including engineering and financial applications. Inevitably any quantity of interest (QoI) Q, say, that is derived from a probabilistic model's output is a random variable. Consequently, important tasks such as predictions, decision making, or optimization based on the available system knowledge Q need to be carried our subject to uncertainties. To reliably account for the effects of these uncertainties, it is indispensable to characterize the distribution of Q or some statistics of it. While the accompanying Part II of this work will be devoted to approximations to the cumulative distribution function (CDF) and robustness indicators such as quantiles (also known as value-at-risk) and the conditional value-at-risk, here we focus on the commonly used approach of characterizing a random system output's distribution by a carefully chosen selection of moments of Q. In fact, in addition to the mean, many important features of distribution, such as location, dispersion, or asymmetry, can be assessed through moments of the deviation from the mean. These statistical moments that are computed about the mean

*Corresponding author

Email addresses: `michele.pisaroni@epfl.ch` (M. Pisaroni), `sebastian.krumscheid@epfl.ch` (S. Krumscheid), `fabio.nobile@epfl.ch` (F. Nobile)

[1] The `cmlmc-py` python library is available at `https://bitbucket.org/mpisaroni_uq/cmlmc-py`

are called central moments. Specifically, the p-th central moment μ_p of a random variable Q is defined as

$$\mu_p = \mathbb{E}\big[(Q-\mu)^p\big] \,, \quad \text{with} \quad \mu = \mathbb{E}[Q] \,,$$

provided the right-hand side exists. Following the definition of μ_p, the first central moment (i.e. $p=1$) is equal to zero. The second central moment μ_2 is the variance (often also denoted as σ^2) and, together with the mean $\mu = \mathbb{E}[Q]$, is one of the most commonly used quantities in applications to characterize a random variable.

The third central moment μ_3 offers insight into the asymmetry of a random variable's distribution about its mean. Specifically, the skewness $\gamma = \mu_3/\sqrt{\mu_2^3}$, which is the standardized counterpart of the third central moment, is commonly used as a measure of probability distribution's asymmetry. Indeed, the skewness (or equivalently the third central moment μ_3) of a symmetric distribution about the mean is zero. Negative values of the skewness indicate that the probability distribution has a left tail that is longer compared to the right one. Analogously, positive values indicate a longer right tail. A measure of a probability distribution's asymmetry is very important in many engineering risk/reliability assessments and financial applications related to stock prices and assets. In fact, Mandelbrot et al. [14] observed that the majority of financial assets returns are non-normal. This is due to the appearance of extreme events more likely than predicted by a normal distribution [4] and due to the fact that crashes occur more often than booms [17]. For this reason, investment decisions based only on the mean and variance cannot discriminate whether a given future event will be more or less likely to appear on the left or right side of the mean [13, 15]. Applied to investment returns, negatively skewed distributions indicate greater chance of extremely negative outcomes, while in positive skewed distributions extremely bad scenarios are not as likely. Assuming a normal distribution, when in fact data sets are skewed, can lead to the so called skewness risk [5]. Similar problems arise in many applications across science and technology where decisions based on a reliability or risk measure need to be taken.

The fourth central moment μ_4 and its standardized counterpart, which is known as kurtosis $Kurt = \mu_4/\mu_2^2$, provide some further important insights into a random variable's distribution. In fact, the kurtosis can be used to measure whether the output random variable are heavy-tailed (high level of kurtosis) or light-tailed (low level of kurtosis) compared to a normal distribution, for which $Kurt = 3$. Heavy-tailed distributions are common in problems where extreme events are likely to appear. Random variables with low levels of kurtosis tend to have light tails and lack of extreme events. In other words high levels of kurtosis indicate that most of the variability in the distribution is due to extreme deviations from the mean.

In this work we consider efficient sampling-based estimators for central moments of a quantity of interest Q output of a complex probabilistic model (such as a fluid flow with random inflow conditions, a stochastic dynamical system, a random partial differential equation, etc.). We address in particular probabilistic models that involve differential equations (such as system of SDEs, SPDEs, or PDEs with random input parameters) for which typically the random system output Q cannot be sampled exactly and only approximate sampling can be accessed with a given accuracy (e.g. by solving the differential equation via some numerical scheme). As a consequence of this inexact sampling, a bias is introduced that has to be accounted for. In the context of estimating the mean $\mathbb{E}[Q]$ of Q, the multilevel Monte Carlo (MLMC) method has established itself as a versatile and efficient sampling-based method

when only inexact sampling is possible [6, 9, 10, 19]. Specifically, let Q_M denote an approximation of the unknown random variable Q obtained via a numerical scheme with M degrees of freedom, i.e. via a numerical scheme that reduces the original infinite dimensional problem to a M-dimensional one and has a computational cost proportional to some power of M. The key idea of the MLMC method then is to consider not just approximations obtained from said scheme with M degrees of freedom, but also approximations obtained with different numbers of degrees of freedoms. In fact, a hierarchy of $L+1$ approximation levels is considered, where each approximation level is defined by its size and the different levels are related by $M_0 < M_1 < \cdots < M_L$. Using N_ℓ independent and identically distributed (i.i.d.) realizations $\omega_{i,\ell}$, $i = 1, \cdots, N_\ell$, of the system's random input parameter on each level $0 \leq \ell \leq L$, the MLMC estimator for the mean $\mu = \mathbb{E}[Q]$ of Q is

$$\mathtt{E}^{\mathtt{MC}}\big(\boldsymbol{Q}^0_{N_0,M_0}\big) + \sum_{\ell=1}^{L} \Big(\mathtt{E}^{\mathtt{MC}}\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big) - \mathtt{E}^{\mathtt{MC}}\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big) \Big) \ , \tag{1}$$

where $\mathtt{E}^{\mathtt{MC}}$ denotes the sample average operator and $\boldsymbol{Q}^\ell_{N_\ell,M_\ell} := \big(Q_{M_\ell}(\omega_{i,\ell})\big)_{i=1,\ldots,N_\ell}$ is the sample of N_ℓ i.i.d. replications of Q_{M_ℓ}. Here, the same $\omega_{i,\ell}$ are used in both $\boldsymbol{Q}_{N_\ell,M_\ell}$ and $\boldsymbol{Q}_{N_\ell,M_{\ell-1}}$, which yields a strong correlation and, hopefully, an estimator with smaller variance.

In this work we extend the MLMC concept to the estimation of arbitrary order central moments μ_p. Specifically, we introduce and analyze a novel multilevel Monte Carlo method that allows an efficient sampling-based estimation from inexact/approximate samples. One of the method's key ingredients is the use of h-statistics [8] as unbiased central moment estimators with minimal variance for the level-wise contributions. That is, instead of the level-wise contributions $\mathtt{E}^{\mathtt{MC}}\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big) - \mathtt{E}^{\mathtt{MC}}\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big)$ that are used in the estimation of the mean (c.f. (1)), here we use terms of the form $h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big) - h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big)$, where h_p denotes an appropriate h-statistic of order p. Consequently, the MLMC estimator $\mathtt{m}_p^{\mathtt{MLMC}}$ for arbitrary order p central moment considered here is of the form

$$\mathtt{m}_p^{\mathtt{MLMC}} = h_p\big(\boldsymbol{Q}^0_{N_0,M_0}\big) + \sum_{\ell=1}^{L} \Big(h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big) - h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big) \Big) \ .$$

We note that a multilevel Monte Carlo estimator for the variance μ_2 of a random variable Q has already been introduced in [2]. There, the authors define the multilevel Monte Carlo estimator by telescoping on the unbiased sample variance estimator for the level-wise contributions. Our approach based on h-statistics thus offers an alternative derivation of said estimator, which allows for a more straightforward complexity analysis in fact. Moreover, the approach introduced here is easily generalized to arbitrary order central moments, as we will illustrate in the following. In fact, the results presented here for estimating μ_p for $p \geq 3$ appear to be novel. Finally, we mention that somewhat related work on multilevel Monte Carlo techniques for arbitrary order central moment estimators can be found in [3]. However, there the authors construct the estimators for $p \geq 3$ based on biased estimators for the level-wise contributions. Consequently, the method introduced in the aforementioned work requires to carefully control this additional bias. Moreover, the mean squared error analysis is also affected by this bias, in the sense that the error is quantified using worst-case bounds based on triangle inequalities. Instead our work, as mentioned earlier already,

uses h-statistics for level-wise contributions, which are unbiased estimators with minimal variance. In fact, these unbiased estimators can be straightforwardly derived in closed-form, allowing for a possibly sharper mean squared error bound. The cost of working with unbiased estimators is that deriving these estimators in closed-form requires somewhat tedious calculations. However, these calculations can be easily carried out automatically by symbolic computer algebra systems, such as `Maple` and `Mathematica`, as we will describe in the following. Lastly, we present a complete algorithm and detail how to tune the MLMC method for central moments to achieve optimal complexity. In fact, we demonstrate the effectiveness of the developed methodology, which has been implemented in our Python library `cmlmc-py`, by applying it to a number of selected examples. These examples include a relevant problem in compressible aerodynamics, namely a transonic airfoil affected by both operating and geometric uncertainties, which we analyze based on estimates of the first four statistical central moments.

The rest of this work is organized as follows. In Sects. 2–3 we present and analyze the sampling-based estimation of central moments. Specifically, in Sect. 2 we first consider the classic Monte Carlo method, before introducing the novel multilevel Monte Carlo estimator in Sect. 3. Following these theoretical considerations, we discuss various practical aspects and implementation details for the multilevel Monte Carlo methods in Sect. 4. In Sect. 5 we apply the MLMC methodology to the above mentioned numerical examples. Finally, we offer a summary and a discussion of our results in Sect. 6.

2. Monte Carlo estimation of central moments

The p-th central moment $\mu_p \equiv \mu_p(Q)$ of a random variable Q (also known as the p-th moment about the mean) is given by

$$\mu_p(Q) := \mathbb{E}\Big[\big(Q - \mathbb{E}[Q]\big)^p\Big] ,$$

for any $p \in \mathbb{N}$ provided it exists, although the value for $p = 1$ is trivial ($\mu_1(Q) = 0$). Any central moment can, of course, be expressed in terms of non-centered (so-called raw moments or moments about the origin) as a consequence of the binomial theorem and the linearity of the expected value:

$$\mu_p(Q) \equiv \mathbb{E}\Big[\big(Q - \mathbb{E}[Q]\big)^p\Big] = \sum_{j=0}^{n} \binom{p}{j} (-1)^{p-j} \mathbb{E}\big[Q^j\big] \mathbb{E}[Q]^{p-j} .$$

However, approximating the p-th central moment $\mu_p(Q)$ by a combination of approximated non-centered moments can be numerically unstable. This may be especially severe if the central moments are small whereas the raw moments are not. To avoid these numerical instabilities, here we present Monte Carlo sampling based estimators for central moments directly. We begin by reviewing classic (single-level) Monte Carlo estimators in this Section, before addressing the multilevel estimators in Sect. 3.

Starting point for the construction of efficient sampling based estimators for central moments are the so-called h-statistics [8]. That is, in the classic single-level setting we consider an i.i.d. sample $\boldsymbol{Q}_N := \big(Q(\omega_i)\big)_{i=1,\dots,N}$, of size N, where each $Q(\omega_i)$ has the same

distribution as Q. The h-statistic $h_p \equiv h_p(\boldsymbol{Q}_N)$ then is an unbiased estimator of $\mu_p(Q)$, in the sense that $\mathbb{E}\big[h_p(\boldsymbol{Q}_N)\big] = \mu_p(Q)$. Moreover, the h-statistic has the favorable property that its variance $\mathbb{V}\mathrm{ar}\big[h_p(\boldsymbol{Q}_N)\big] = \mathbb{E}\big[\big(h_p(\boldsymbol{Q}_N) - \mu_p(Q)\big)^2\big]$ is minimal compared to all other unbiased estimators [12]. Based on the sample $\boldsymbol{Q}_N$ of size N, the h-statistic $h_p(Q)$ is commonly expressed in terms of power sums $S_a \equiv S_a(\boldsymbol{Q}_N) := \sum_{i=1}^N Q(\omega_i)^a$. For example, the first three h-statistics are

$$
\begin{aligned}
h_2 &= \frac{NS_2 - S_1^2}{(N-1)N} \,, \\
h_3 &= \frac{N^2S_3 - 3NS_2S_1 + 2S_1^3}{(N-2)(N-1)N} \,, \\
h_4 &= \frac{(-4N^2+8N-12)\,S_3S_1 + (N^3-2N^2+3N)\,S_4 + 6NS_2S_1^2 + (9-6N)S_2^2 - 3S_1^4}{(N-3)(N-2)(N-1)N} \,,
\end{aligned}
$$

where we have used the shorthand notation $h_p \equiv h_p(\boldsymbol{Q}_N)$ and $S_a \equiv S_a(\boldsymbol{Q}_N)$ for brevity (see, e.g., [8] for the construction of h_p for arbitrary p).

In practice, sampling the random variable Q usually requires the solution of a complex problem (e.g. fluid flow with random initial/boundary conditions, random dynamical system, etc.), which inevitably involves a discretization step. That is, it is often not possible to sample Q exactly and instead we assume that one can only draw approximate i.i.d. random variables $Q_M(\omega_i)$, $i = 1, \dots, N$, from a random variable Q_M, which is a suitable approximation (in a sense made precise below) of the unknown random variable Q. In this case the natural Monte Carlo (MC) estimator for the p-th central moment $\mu_p(Q)$ by means of an i.i.d. sample $\boldsymbol{Q}_{N,M} := \big(Q_M(\omega_i)\big)_{i=1,\dots,N}$ of the approximate, computable random variable $\boldsymbol{Q}_M$ is simply the h-statistic based on $\boldsymbol{Q}_{N,M}$:

$$
\mathtt{m}_p^{\mathtt{MC}} := h_p(\boldsymbol{Q}_{N,M}) \,.
$$

That is, there are two levels of approximations: the first one due to approximate sampling ($\mu_p(Q) \approx \mu_p(Q_M)$) and the second one due to the Monte Carlo error ($\mu_p(Q_M) \approx \mathtt{m}_p^{\mathtt{MC}}$). Consequently, the mean squared error of this Monte Carlo estimator is

$$
\mathrm{MSE}\big(\mathtt{m}_p^{\mathtt{MC}}\big) := \mathbb{E}\Big[\big(\mathtt{m}_p^{\mathtt{MC}} - \mu_p(Q)\big)^2\Big] = \big(\mu_p(Q_M) - \mu_p(Q)\big)^2 + \mathbb{V}ar\big[h_p(\boldsymbol{Q}_{N,M})\big] \,, \tag{2}
$$

from which we identify the bias $\big|\mu_p(Q_M) - \mu_p(Q)\big|$ and the statistical error $\mathbb{V}\mathrm{ar}\big(h_p(\boldsymbol{Q}_{N,M})\big)$. Under appropriate assumptions, the statistical error is of order $\mathcal{O}(N^{-1})$ as usual. In fact, for the first three central moment estimators ($h_1 \equiv 0$ not included), the MC estimator's variance reads

$$
\mathbb{V}ar\big(h_2\big) = \frac{\mu_4}{N} - \frac{\mu_2^2(N-3)}{(N-1)N} \,, \tag{3a}
$$

$$
\mathbb{V}ar\big(h_3\big) = \frac{3\mu_2^3\,(3N^2-12N+20)}{(N-2)(N-1)N} - \frac{3\mu_4\mu_2(2N-5)}{(N-1)N} + \frac{\mu_6}{N} - \frac{\mu_3^2(N-10)}{(N-1)N} \,, \tag{3b}
$$

$$
\begin{aligned}
\mathbb{V}ar\big(h_4\big) = {} & \frac{72\mu_2^4\,(N^2-6N+12)}{(N-3)(N-2)(N-1)N} + \frac{16\mu_3^2\mu_2\,(N^2-4N+13)}{(N-2)(N-1)N} \\
& - \frac{24\mu_4\mu_2^2(4N-11)}{(N-2)(N-1)N} + \frac{16\mu_6\mu_2}{(N-1)N} + \frac{\mu_8}{N} - \frac{8\mu_3\mu_5}{N} - \frac{\mu_4^2(N-17)}{(N-1)N} \,,
\end{aligned} \tag{3c}
$$

where we have suppressed the arguments of $h_p \equiv h_p(\boldsymbol{Q}_{N,M})$ and $\mu_p \equiv \mu_p(Q_M)$ for brevity again. It is noteworthy that these quantities can be computed (combinatorial problem) straightforwardly for any p using the `Mathematica` package `mathstatica` [16], due to the h-statistic's power sum representation.

If one assumes that the approximate random variable Q_M is such that the bias term $\left|\mu_p(Q_M) - \mu_p(Q)\right|$ decays at a certain rate when increasing the discretization parameter M, then it is possible to balance the squared bias and statistical error contributions to the MSE in (2). Such a bias assumption is plausible since the bias term is related to the numerical method (assumed to be consistent) used to approximate the underlying complex system. At the same time, generating realizations of Q_M typically becomes more expensive as M increases. The following result thus quantifies the computational cost to estimate the p-th central moment by the MC method, when using optimal discretization parameter M and optimal sample size N to achieve a prescribed accuracy. As a matter of fact, the theoretical result below is the central moment analog of the standard result for expectations.

Proposition 2.1. *Let $p \in \mathbb{N}$, $p \geq 2$, and assume that the $2p$-th central moment of Q_M is bounded, so that $\mu_{2p}(Q_M) < \infty$ for $M \gg 1$. Furthermore, suppose that there exist constants α and γ such that*

(i) the bias decays with order $\alpha > 0$, in the sense that $\left|\mu_p(Q_M) - \mu_p(Q)\right| \leq c_\alpha M^{-\alpha}$ for some constant $c_\alpha > 0$,

(ii) the cost to compute each i.i.d. realization of Q_M is bounded by $\mathrm{cost}(Q_M) \leq c_\gamma M^\gamma$ for some constants $c_\gamma, \gamma > 0$.

The MC estimator $\mathtt{m}_p^{\mathtt{MC}} = h_p(\boldsymbol{Q}_{N,M})$ with $N = \mathcal{O}(\varepsilon^{-2})$ and $M = \mathcal{O}(\varepsilon^{-1/\alpha})$ satisfies $\mathrm{MSE}\big(\mathtt{m}_p^{\mathtt{MC}}\big) = \mathcal{O}(\varepsilon^2)$ and the cost associated with computing this estimator is bounded by

$$\mathrm{cost}\big(\mathtt{m}_p^{\mathtt{MC}}\big) = N \cdot \mathrm{cost}(Q_M) \leq c\varepsilon^{-2-\gamma/\alpha} \,,$$

where c is independent of $\varepsilon > 0$.

Note that the constants appearing in the result above (i.e. c, c_α, and c_γ) depend on the order p of the central moment. In fact, also the rates α and γ may depend on p in principle. However, numerical evidence suggests that the rates may, in fact, not depend on p for a large class of problems; cf. the numerical studies presented in Sect. 5.

2.1. Practical aspect: MSE and unbiased variance estimation

A robust implementation of the MC estimator $\mathtt{m}_p^{\mathtt{MC}}$ should also provide an estimation of the associated MSE. This is also the first step towards building an adaptive MC algorithm in which the sample size N and/or the discretization parameter M are progressively increased to achieve a MSE smaller than a prescribed tolerance.

The bias term $\left|\mu_p(Q_M) - \mu_p(Q)\right|$ relates only to the numerical discretization of the underlying differential problem. Possible ways of estimating the bias include:

(i) the calculation on a sequence of refined discretizations with parameters $M_1 < M_2 < \ldots$ and extrapolation of the error;

$\frac{\hat{V}_2}{N}$	$\frac{N\Big((N-1)^2NS_4-(N^2-3)S_2^2\Big)+(6-4N)S_1^4+4N(2N-3)S_2S_1^2-4(N-1)^2NS_3S_1}{(N-3)(N-2)(N-1)^2N^2}$
$\frac{\hat{V}_3}{N}$	$\frac{1}{(N-5)(N-4)(N-3)(N-2)^2(N-1)^2N^2}\Big(-12(3N^2-15N+20)S_1^6$ $+36N(3N^2-15N+20)S_2S_1^4-24N^2(2N^2-9N+11)S_3S_1^3$ $+3NS_1^2\big((7N^4-36N^3+79N^2-90N+40)S_4-6N(4N^2-21N+29)S_2^2\big)$ $-6NS_1\big((N^3-3N^2+6N-8)(N-1)^2S_5+(-5N^4+18N^3+13N^2-90N+40)S_2S_3\big)$ $+N\big((N-1)^2N(N^3-3N^2+6N-8)S_6+3(3N^4-24N^3+71N^2-90N+40)S_2^3$ $-3(2N^5-11N^4+14N^3+25N^2-70N+40)S_4S_2$ $-(N^5+4N^4-41N^3+40N^2+100N-80)S_3^2\big)\Big)$

Table 1: Closed-form expressions of the unbiased estimators $\hat{V}_p/N$ for $\mathbb{V}\mathrm{ar}\big[h_p(\boldsymbol{Q}_{N,M})\big] = V_p/N$, $p = 2, 3$, as polynomial functions of the power sums $S_a \equiv S_a(\boldsymbol{Q}_{N,M})$.

(ii) error estimations based on a-posteriori error estimators (see e.g. [21], [1]) available for certain type of equations.

We will not detail further this aspect here, as the main goal of this work is on the estimation of the statistical error. For this, a possibly unbiased estimator for the variance $\mathbb{V}ar\big[h_p(\boldsymbol{Q}_{N,M})\big]$ based on the same sample $\boldsymbol{Q}_{N,M}$ of size N is needed. We discuss hereafter the derivation of one such estimator. As we have seen in (3), it holds that $\mathbb{V}ar\big[h_p(\boldsymbol{Q}_{N,M})\big] = \mathcal{O}(1/N)$. It is thus convenient to set $V_p := N \cdot \mathbb{V}ar\big[h_p(\boldsymbol{Q}_{N,M})\big]$ and derive unbiased estimators $\hat{V}_p$ of V_p. However, the naive approach of simply replacing μ_k, for $k = 2, \ldots, 2p$, in (3) by its unbiased estimator h_k will not result in an unbiased estimator for V_p, since the statistical error $\mathbb{V}ar\big[h_p(\boldsymbol{Q}_{N,M})\big]$ depends non-linearly on the central moments. Instead, we do not only substitute h_k for μ_k but also introduce an additional multiplicative coefficient for each substitution. For example, inspecting equation (3a) suggests to make the ansatz $\hat{V}_2 = a_1h_4 + a_2h_2^2$ for $p = 2$. Similarly, (3b) implies the ansatz $\hat{V}_3 = a_1h_2^3 + a_2h_2h_4 + a_3h_6 + a_4h_3^2$ for the case $p = 3$ and so on. For an ansatz of this form the expected value of $\hat{V}_p$, $\mathbb{E}[\hat{V}_p]$, can be computed as a polynomial function of the central moments μ_k, $k = 2, \ldots, 2p$, using `mathstatica`. Consequently, we can derive unbiased estimators by equating the coefficients of such polynomial with the corresponding ones in the expression of $\mathbb{E}[V_p]$.For example, for $p = 2$ we find

$$\mathbb{E}[\hat{V}_2] = \frac{\mu_4(a_2 + a_1N)}{N} + \frac{a_2\mu_2^2\,(N^2 - 2N + 3)}{(N-1)N}\,,$$

which, after comparing with equation (3a), yields $a_1 = \frac{N-1}{N^2-2N+3}$ and $a_2 = -\frac{N-3}{N^2-2N+3}$. The unbiased variance estimators $\hat{V}_p/N$ of $\mathbb{V}\mathrm{ar}\big[h_p(\boldsymbol{Q}_{N,M})\big]$ obtained by following this procedure are summarized in Table 1, where the final expression has been given directly in terms of the power sums $S_a \equiv S_a(\boldsymbol{Q}_{N,M})$ instead of the h-statistics. For the sake of a clear presentation, we present the unbiased estimator for the case $p = 4$ in AppendixA. It is noteworthy, that although these formulas are rather lengthy, they are in closed-form, so that they are easily implementable and are available in the accompanying Python library `cmlmc-py`.

3. Multilevel Monte Carlo estimation of central moments.

Using the results presented in the previous Section and following the general construction of MLMC estimators, we introduce the MLMC estimator for the p-th central moment as:

$$\mathtt{m}_p^{\mathtt{MLMC}} := h_p\big(\boldsymbol{Q}^0_{N_0,M_0}\big)+\sum_{\ell=1}^{L}\Big(h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big)-h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big)\Big)\equiv\sum_{\ell=0}^{L}\Big(h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big)-h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big)\Big)\ , \tag{4}$$

with the convention that $h_p\big(\boldsymbol{Q}^0_{N_0,M_{-1}}\big)\equiv 0$. Here, the sample $\boldsymbol{Q}^\ell_{N_\ell,M_\ell}$ of i.i.d. realizations is given by $\boldsymbol{Q}^\ell_{N_\ell,M_\ell} := \big(Q_{M_\ell}(\omega_{i,\ell})\big)_{i=1,\dots,N_\ell}$ for any level ℓ. The superscript ℓ of both samples $\boldsymbol{Q}^\ell_{N_\ell,M_\ell}$ and $\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}$ is used to indicate the correlation across two consecutive levels, which is the key ingredient for any multilevel Monte Carlo method. Specifically, the realizations of the sample $\boldsymbol{Q}^\ell_{\cdot,\cdot}$ and those of $\boldsymbol{Q}^k_{\cdot,\cdot}$, $k\neq l$, are independent, while the N_ℓ realizations of $\boldsymbol{Q}^\ell_{N_\ell,M_\ell}$ and $\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}$ are correlated, in the sense that the approximate quantities of interest computed on the finer discretization (i.e. sample $\boldsymbol{Q}_{N_\ell,M_\ell}$) and those computed on the coarser discretization (i.e. samples of $\boldsymbol{Q}_{N_\ell,M_{\ell-1}}$) correspond to the same realizations of the uncertain inputs. Consequently, the MLMC estimator's mean squared error is

$$\mathrm{MSE}\big(\mathtt{m}_\mathtt{p}^{\mathtt{MLMC}}\big)=\big(\mu_p(Q_{M_L})-\mu_p(Q)\big)^2+\sum_{\ell=0}^{L}\mathbb{V}ar\big[\Delta_\ell h_p\big]\ , \tag{5}$$

where we have introduced the shorthand notation

$$\Delta_\ell h_p\equiv\Delta_\ell h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell},\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big):=h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big)-h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big)\ .$$

The bias term $|\mu_p(Q_{M_L})-\mu_p(Q)|$ in (5) corresponds to the bias of the classic Monte Carlo method described in Sect. 2 on discretization level L, cf. equation (2). The analysis of the variances $\mathbb{V}\mathrm{ar}[\Delta_\ell h_p]$ and their dependence on N_ℓ as well as on the central moments of $\boldsymbol{Q}^\ell_{N_\ell,M_\ell}$ and $\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}$ is more cumbersome than for the classic MC estimator. In particular we need to quantify the correlation between $\boldsymbol{Q}^\ell_{N_\ell,M_\ell}$ and $\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}$. To do so, it is convenient to introduce both the sample sum and the sample difference of these samples:

$$\begin{aligned}\boldsymbol{X}^{\ell,+}_{N_\ell}&:=\big(X^{\ell,+}_i\big)_{i=1,\dots,N_\ell}\quad\text{with}\quad X^{\ell,+}_i:=Q_{M_\ell}(\omega_{i,\ell})+Q_{M_{\ell-1}}(\omega_{i,\ell})\ ,\\ \boldsymbol{X}^{\ell,-}_{N_\ell}&:=\big(X^{\ell,-}_i\big)_{i=1,\dots,N_\ell}\quad\text{with}\quad X^{\ell,-}_i:=Q_{M_\ell}(\omega_{i,\ell})-Q_{M_{\ell-1}}(\omega_{i,\ell})\ .\end{aligned}$$

In other words, we have that $\boldsymbol{X}^{\ell,+}_{N_\ell}=\boldsymbol{Q}^\ell_{N_\ell,M_\ell}+\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}$ and $\boldsymbol{X}^{\ell,-}_{N_\ell}=\boldsymbol{Q}^\ell_{N_\ell,M_\ell}-\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}$. Moreover, we introduce the bivariate power sums $S_{a,b}$ analogously to the power sums S_a in the previous Section, that is

$$S_{a,b}\big((X_i)_{i=1,\dots,N},(Y_i)_{i=1,\dots,N}\big):=\sum_{i=1}^{N}X_i{}^aY_i^b\ ,$$

for any two samples $(X_i)_{i=1,\dots,N}$ and $(Y_i)_{i=1,\dots,N}$ of the same size N. Then we can compute the variance $\mathbb{V}\mathrm{ar}\big(\Delta_\ell h_p\big)$ for each level ℓ as follows:

1. For each ℓ, we express the term $\Delta_\ell h_p \equiv h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big) - h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big)$ in terms of bivariate power series $S_{a,b}$ in $\boldsymbol{X}^{\ell,+}_{N_\ell}$ and $\boldsymbol{X}^{\ell,-}_{N_\ell}$, that is in terms of

$$S^\ell_{a,b} \equiv S_{a,b}\Big(\boldsymbol{X}^{\ell,+}_{N_\ell}, \boldsymbol{X}^{\ell,-}_{N_\ell}\Big) \,.$$

   This can, of course, be achieved by using the identities $\boldsymbol{Q}^\ell_{N_\ell,M_\ell} = \frac{1}{2}\big(\boldsymbol{X}^{\ell,+}_{N_\ell} + \boldsymbol{X}^{\ell,-}_{N_\ell}\big)$ and $\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}} = \frac{1}{2}\big(\boldsymbol{X}^{\ell,+}_{N_\ell} - \boldsymbol{X}^{\ell,-}_{N_\ell}\big)$ and some algebra.

2. The obtained representation of $\Delta_\ell h_p$ in terms of these bivariate power sums in $\boldsymbol{X}^{\ell,+}_{N_\ell}$ and $\boldsymbol{X}^{\ell,-}_{N_\ell}$ is then amenable for further treatment by the `mathstatica` software. In fact, the software provides an efficient algorithm for treating the combinatorial problem of computing the desired variances, due to the power series representation.

Following this procedure, the first step yields for example

$$\Delta_\ell h_2 = \frac{N_\ell S^\ell_{1,1} - S^\ell_{0,1} S^\ell_{1,0}}{(N_\ell - 1) N_\ell} \,,$$
$$\Delta_\ell h_3 = -\frac{-{N_\ell}^2 S^\ell_{0,3} - 3{N_\ell}^2 S^\ell_{2,1} + 3N_\ell S^\ell_{0,2} S^\ell_{0,1} + 3N_\ell S^\ell_{2,0} S^\ell_{0,1} + 6N_\ell S^\ell_{1,0} S^\ell_{1,1} - 2S^{\ell 3}_{0,1} - 6S^{\ell 2}_{1,0} S^\ell_{0,1}}{4(N_\ell - 2)(N_\ell - 1)N_\ell} \,,$$

where we have again omitted the arguments for brevity. The same procedure can also be used to derive close-form expressions for $\Delta_\ell h_p$ with $p \geq 4$, which become rather lengthy and are thus not presented here for the sake of a clear presentation. Based on these closed-form expressions for $\Delta_\ell h_p$, the required expression of the variance $\mathbb{V}ar\big[\Delta_\ell h_p\big]$ on level ℓ then follows accordingly as

$$\mathbb{V}\mathrm{ar}\big[\Delta_\ell h_2\big] = -\frac{(N_\ell - 2)\mu^2_{1,1}}{(N_\ell - 1)N_\ell} + \frac{\mu_{0,2}\mu_{2,0}}{(N_\ell - 1)N_\ell} + \frac{\mu_{2,2}}{N_\ell} \,, \tag{6a}$$

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\big[\Delta_\ell h_3\big] = {} & \frac{3\left(3{N_\ell}^2 - 12N_\ell + 20\right)\mu^3_{0,2}}{16(N_\ell - 2)(N_\ell - 1)N_\ell} + \frac{9\left({N_\ell}^2 - 4N_\ell + 8\right)\mu^2_{1,1}\mu_{0,2}}{4(N_\ell - 2)(N_\ell - 1)N_\ell} \\
& + \frac{9\left({N_\ell}^2 - 4N_\ell + 12\right)\mu^2_{2,0}\mu_{0,2}}{16(N_\ell - 2)(N_\ell - 1)N_\ell} + \frac{9\left({N_\ell}^2 - 4N_\ell + 6\right)\mu^2_{1,1}\mu_{2,0}}{2(N_\ell - 2)(N_\ell - 1)N_\ell} \\
& - + \frac{9(N_\ell - 2)\mu_{2,0}\mu^2_{0,2}}{8(N_\ell - 1)N_\ell} + \frac{9\mu_{4,0}\mu_{0,2}}{16(N_\ell - 1)N_\ell} - \frac{9(N_\ell - 2)\mu_{2,2}\mu_{0,2}}{8(N_\ell - 1)N_\ell} \\
& - \frac{3(2N_\ell - 5)\mu_{0,4}\mu_{0,2}}{16(N_\ell - 1)N_\ell} + \frac{9\mu^2_{1,2}}{4(N_\ell - 1)N_\ell} + \frac{\mu_{0,6}}{16N_\ell} + \frac{3\mu_{2,4}}{8N_\ell} \\
& + \frac{9\mu_{1,2}\mu_{3,0}}{4(N_\ell - 1)N_\ell} + \frac{9\mu_{4,2}}{16N_\ell} - \frac{3\mu_{0,4}\mu_{2,0}}{8N_\ell} - \frac{3(N_\ell - 4)\mu_{1,1}\mu_{1,3}}{4(N_\ell - 1)N_\ell} \\
& - \frac{9(N_\ell - 2)\mu_{1,1}\mu_{3,1}}{4(N_\ell - 1)N_\ell} - \frac{3(N_\ell - 4)\mu_{0,3}\mu_{2,1}}{8(N_\ell - 1)N_\ell} - \frac{9(N_\ell - 3)\mu_{2,0}\mu_{2,2}}{8(N_\ell - 1)N_\ell} \\
& - \frac{(N_\ell - 10)\mu^2_{0,3}}{16(N_\ell - 1)N_\ell} - \frac{9(N_\ell - 6)\mu^2_{2,1}}{16(N_\ell - 1)N_\ell} \,,
\end{aligned} \tag{6b}$$

where we present $\mathbb{V}\mathrm{ar}\big[\Delta_\ell h_4\big]$ in AppendixA for a clearer presentation. Here, $\mu_{p,q} \equiv \mu_{p,q}\big(X^{\ell,+}, X^{\ell,-}\big)$ denotes the bivariate central moment of order (p,q) of $X^{\ell,+}$ and $X^{\ell,-}$, where the bivariate central moment is given by

$$\mu_{p,q}(X,Y) := \mathbb{E}\Big[\big(X - \mathbb{E}(X)\big)^p \big(Y - \mathbb{E}(Y)\big)^q\Big] ,$$

for any two random variables X and Y.

Inspection of the variance expressions for $\mathbb{V}\mathrm{ar}\big(\Delta_\ell h_p\big)$ in (6) reveals that $\mathbb{V}\mathrm{ar}\big(\Delta_\ell h_p\big) = \mathcal{O}(1/N_\ell)$ for any fixed ℓ. Setting $V_{\ell,p} := N_\ell \mathbb{V}\mathrm{ar}\big(\Delta_\ell h_p\big)$, the mean squared error of the MLMC estimator $\mathtt{m}_{\mathtt{p}}^{\mathtt{MLMC}}$ can then be written in the somewhat more familiar form

$$\mathrm{MSE}\big(\mathtt{m}_p^{\mathtt{MLMC}}\big) = \big(\mu_p(Q_{M_L}) - \mu_p(Q)\big)^2 + \sum_{\ell=0}^{L} \frac{V_{\ell,p}}{N_\ell} ,$$

which indicates the usual interplay of bias and statistical error. Due to the identities for the variance expressions, the complexity result for the MLMC estimator for central moments follows by the same arguments as the ones used in the standard MLMC result; see, e.g., [10]. In fact, the only difference to the standard MLMC complexity result is that the notion of bias and variance have to be modified. Then even the formulas for the optimal number of levels and sample size on each level follow immediately; see Sect. 4 for further details.

Proposition 3.1. *Let $p \in \mathbb{N}$, $p \geq 2$, and assume that the $2p$-th central moment of Q_{M_ℓ} is bounded, so that $\mu_{2p}(Q_{M_\ell}) < \infty$, for $\ell \geq 0$. Furthermore, suppose that there exist constants α, β, and γ such that $2\alpha \geq \min(\beta, \gamma)$ and*

(i) the bias decays with order $\alpha > 0$, in the sense that $\big|\mu_p(Q_{M_\ell}) - \mu_p(Q)\big| \leq c_\alpha M_\ell^{-\alpha}$ for some constant $c_\alpha > 0$,

(ii) the variance $V_{\ell,p} \equiv \mathbb{V}ar\big[\Delta_\ell h_p\big] N_\ell$ decays with order $\beta > 0$, in the sense that $V_{\ell,p} \leq c_\beta M_\ell^{-\beta}$ for some constant $c_\beta > 0$,

(iii) the cost to compute each i.i.d. realization of Q_{M_ℓ} is bounded by $\mathrm{cost}(Q_{M_\ell}) \leq c_\gamma M_\ell^{\gamma}$ for some constants $c_\gamma, \gamma > 0$.

For any $0 < \varepsilon < e^{-1}$, the MLMC estimator $\mathtt{m}_p^{\mathtt{MLMC}}$ with maximum level $L \in \mathbb{N}_0$ such that $\big|\mu_p(Q_{M_L}) - \mu_p(Q)\big| \leq \frac{\varepsilon}{\sqrt{2}}$ and with sample size $N_\ell \in \mathbb{N}$ on level ℓ given by

$$N_\ell = \left\lceil \frac{2}{\varepsilon^2} \sqrt{\frac{V_{\ell,p}}{\mathrm{cost}(Q_{M_\ell})}} \sum_{j=0}^{L} \sqrt{\mathrm{cost}(Q_{M_j}) V_{j,p}} \right\rceil , \quad 0 \leq \ell \leq L ,$$

satisfies $\mathrm{MSE}\big(\mathtt{m}_p^{\mathtt{MLMC}}\big) \leq \varepsilon^2$ at a computational cost that is bounded by

$$\mathrm{cost}\big(\mathtt{m}_p^{\mathtt{MLMC}}\big) \leq c \begin{cases} \varepsilon^{-2} \ln(\varepsilon^{-1})^2 , & \text{if } \beta = \gamma , \\ \varepsilon^{-\left(2 + \frac{\gamma - \beta}{\alpha}\right)} , & \text{if } \beta < \gamma , \\ \varepsilon^{-2} , & \text{if } \beta > \gamma , \end{cases}$$

where c is independent of $\varepsilon > 0$.

As remarked after Prop. 2.1 already, it is also the case for the MLMC estimator that the appearing constants depend on the order p. It may also be the case that the rates depend on p, although numerical experiments suggest that this is not the case for a large class of problems; see Sect. 5. Finally, we mention that the proposition above can be stated in terms of the $\mathrm{cost}(Q_{M_\ell})$ instead of the $\mathrm{cost}(\Delta_\ell h_p)$ due to the availability of the closed-form expressions for $\Delta_\ell h_p$, whose evaluation cost is negligible compared to $\mathrm{cost}(Q_{M_\ell})$.

3.1. Practical aspect: MSE and unbiased level-wise variance estimation

As for the classic Monte Carlo method described in Sect. 2, also a robust implementation of the MLMC estimator should provide an estimation of the associated MSE. Moreover, estimations of $\mathbb{V}ar\big[\Delta_\ell h_p\big]$ are further needed to determine the optimal sample size N_ℓ on each level to achieve a prescribed tolerance ε, and we detail hereafter a practical construction of unbiased estimators for $V_{\ell,p}$.

Concerning the bias term, the same considerations made for the classic MC estimator hold here as well. However, since the MLMC estimator already uses a sequence of discretizations, the situation is somewhat simplified as a natural way to estimate the bias is $|\mu_p(Q_{M_L}) - \mu_p(Q)| \approx |\Delta_L h_p(Q^L_{N_L,M_L}, Q^L_{N_L,M_{L-1}})|$.

Next, we discuss how to construct an unbiased estimator of the variance $\mathbb{V}ar\big[\Delta_\ell h_p\big]$ on each level ℓ, or equivalently of $V_{\ell,p} \equiv \mathbb{V}ar\big[\Delta_\ell h_p\big] N_\ell$, based on the samples $Q^\ell_{N_\ell,M_\ell}$ and $Q^\ell_{N_\ell,M_{\ell-1}}$. Similarly to the derivation of an unbiased variance estimator for the MC method (cf. Sect. 2.1), an unbiased estimator of the level-wise variance $V_{\ell,p}$ is not straightforward to construct. In fact, here the situation is even slightly more complicated due to the highly nonlinear combination of the bivariate central moments $\mu_{k,l}$, cf. the expressions in (6). However, also for the bivariate central moments $\mu_{k,l}$ there exist unbiased estimators, namely the $h_{k,l}$-statistic [16]. As a consequence, the procedure to construct unbiased variance estimators described in Sect. 2.1 can be followed for the most parts with only minor modifications. Specifically, to construct unbiased estimators of $\mathbb{V}ar\big[\Delta_\ell h_p\big] \equiv V_{\ell,p}/N_\ell$, we proceed as follows:

1. We make an initial generic ansatz for the estimator $\hat{V}_{\ell,p}$ of $V_{\ell,p}$ based upon replacing the central moments $\mu_{k,l}$ in (6) by their multivariate $h_{k,l}$-statistics, so that $\hat{V}_{\ell,p} = \sum_i a_i h^{m_i}_{p_i,q_i} h^{n_i}_{r_i,s_i}$ with the same powers m_i and n_i appearing in (6).

2. We compute the expectation $\mathbb{E}[\hat{V}_{\ell,p}]$ of the considered ansatz explicitly as a polynomial function of the central moments $\mu_{k,l}$. Again, this combinatorial manipulation can be carried out efficiently using the `mathstatica` software.

3. We assemble a linear system of equations for the unknown coefficients $(a_i)_i$ in the considered ansatz by equating the coefficients in (6) with those of $\mathbb{E}[\hat{V}_{\ell,p}]/N_\ell$, obtained by ordering with respect to the central moments $\mu_{k,l}$.

4. If the linear system is not uniquely solvable, then we augment the ansatz for the estimator to account for the newly introduced central moment terms by computing $\mathbb{E}[\hat{V}_{\ell,p}]$ and repeat steps 2–4.

Obviously, it is also possible to directly consider an ansatz that contains all unique combinations of $\mu^{k_1}_{p_1,q_1}\mu^{k_2}_{p_2,q_2}$, such that $k_1(p_1+q_1) + k_2(p_2+q_2) = 2p$. However, the procedure

described above offers the advantage that it may result in a lower dimensional linear system, which needs to be solved.

We detail here the procedure for $p = 2$. In view of (6a) we first make the initial ansatz $\hat{V}_{\ell,2} = a_1 h_{1,1}^2 + a_2 h_{0,2} h_{2,0} + a_3 h_{2,2}$. Next, we compute the expectation of this ansatz, which can be written as

$$\mathbb{E}[\hat{V}_{\ell,2}] = \frac{2a_2 + \big((N_\ell - 1)^2 + 1\big)\, a_3}{(N_\ell - 1)N_\ell} \mu_{1,1}^2 + \frac{a_2(N_\ell - 1)^2 + a_3}{(N_\ell - 1)N_\ell} \mu_{0,2}\mu_{2,0} + \frac{a_1 N_\ell + a_2 + a_3}{N_\ell} \mu_{2,2} \,.$$

By equating the coefficients of the right-hand side above and those in (6a) we then obtain a linear system of equations for the coefficients a_1, a_2, and a_3. Finally, solving this linear system yields $a_1 = \frac{N_\ell - 1}{{N_\ell}^2 - 2N_\ell + 3}$, $a_2 = \frac{N_\ell - 1}{{N_\ell}^3 - 4{N_\ell}^2 + 7N_\ell - 6}$, and $a_3 = \frac{-{N_\ell}^2 + 4N_\ell - 5}{{N_\ell}^3 - 4{N_\ell}^2 + 7N_\ell - 6}$. Using these coefficients, we can eventually express the unbiased sample-based estimator of $\mathbb{V}ar\big[\Delta_\ell h_2\big] = V_{\ell,2}/N_\ell$ as a polynomial function of the bivariate power sums as

$$\begin{aligned}
\frac{\hat{V}_{\ell,2}}{N_\ell} = \frac{1}{\big(N_\ell - 3\big)\big(N_\ell - 2\big)\big(N_\ell - 1\big)^2 {N_\ell}^2} \Bigg(& N_\ell \Big(\big(-{N_\ell}^2 + N_\ell + 2\big) S_{1,1}^{\ell 2} \\
& + \big(N_\ell - 1\big)^2 \big(N_\ell S_{2,2}^\ell - 2 S_{1,0}^\ell S_{1,2}^\ell\big) + \big(N_\ell - 1\big) S_{0,2}^\ell (S_{1,0}^{\ell 2} - S_{2,0}^\ell) \Big) + S_{0,1}^{\ell 2} \Big(\big(6 - 4N_\ell\big) S_{1,0}^{\ell 2} \\
& + \big(N_\ell - 1\big) N_\ell S_{2,0}^\ell \Big) - 2N_\ell S_{0,1}^\ell \Big(\big(N_\ell - 1\big)^2 S_{2,1}^\ell + \big(5 - 3N_\ell\big) S_{1,0}^\ell S_{1,1}^\ell \Big) \Bigg) \,,
\end{aligned} \tag{7}$$

where $S_{a,b} \equiv S_{a,b}\big(\boldsymbol{X}_{N_\ell}^{\ell,+}, \boldsymbol{X}_{N_\ell}^{\ell,-}\big)$ for brevity. The same procedure can also be applied to obtain unbiased estimators for higher order (i.e. for any $p \geq 2$) central moments, which become rather lengthy though. However, we emphasize that the obtained unbiased variance estimators are in closed-form, so that an efficient implementation is possible. For example, in AppendixB we present the unbiased estimator for the case $p = 3$, while we refer to our implementation details available in the Python library `cmlmc-py` for the formula for $p = 4$.

We reiterate that the procedure introduced here yields unbiased sample-based variance estimators, which are needed for the practical error control and tuning of the MLMC approach introduced in this work; see Section 4 for details. The fact that these variance estimators are unbiased and not just asymptotically unbiased is particularly important on finer levels ℓ, on which the sample size N_ℓ will be small. For example, for $p = 2$ the bias of the naive variance estimator, which is obtained by simply replacing the bivariate central moments $\mu_{k,l}$ by the corresponding $h_{k,l}$-statistics, is $\frac{({N_\ell}^2 - 4N_\ell + 6)\mu_{1,1}{}^2 + (3 - 2N_\ell)\mu_{0,2}\mu_{2,0} - ({N_\ell}^2 - 4N_\ell + 3)\mu_{2,2}}{(N_\ell - 1)^2 {N_\ell}^2}$. Although this additional bias as a function of the sample size N_ℓ is of order $\mathcal{O}({N_\ell}^{-2})$, it may still contribute to a non-negligible error of the MLMC estimator, in particular due to fine levels for which N_ℓ will be small. Finally, we also emphasize that, as a consequence of being based on unbiased estimators, the MLMC method for central moments introduced in this work does not come at the expense of introducing an additional systematic error (i.e. a bias) that needs to be accounted for, unlike other works on central moment estimators, such as [3].

3.2. From mean squared errors to confidence intervals

The discussion of both the MC method and the MLMC method above was solely based on the mean squared error as an accuracy measure. However, for some applications it is often

also desirable to associate confidence intervals (or, equivalently, failure probabilities) to an estimator. Specifically, let $\hat{\theta}$ be a generic estimator of the deterministic value θ with mean squared error given by $\mathrm{MSE}(\hat{\theta}) = \mathbb{E}\big[(\hat{\theta}-\theta)^2\big]$. For a confidence $p_c \in (0,1)$, the associated confidence interval can then be characterized by the value $\delta > 0$, such that

$$\mathbb{P}\Big(\big|\hat{\theta}-\theta\big| < \delta\Big) \geq p_c \,.$$

In the absence of any further knowledge of the probability distribution of $|\hat{\theta}-\theta|$, a sufficient condition for the length δ of the confidence interval can be derived using Chebyshev's inequality:

$$\mathbb{P}\Big(\big|\hat{\theta}-\theta\big| \leq \delta\Big) \leq \frac{\mathrm{MSE}(\hat{\theta})}{\delta^2} = 1-p_c \quad \Rightarrow \quad \delta = \sqrt{\frac{\mathrm{MSE}(\hat{\theta})}{1-p_c}} \,. \tag{8}$$

That is, the confidence interval can be directly linked to the estimator's mean squared error. Consequently, the mean squared error based analysis considered in this work can straightforwardly be used to quantify confidence regions (or failure probabilities) of estimators. It is noteworthy however, that the confidence region identified in (8) may be rather conservative due to the use of Chebyshev's inequality. We will revisit this fact in the forthcoming Part II of this work, where we will introduce methodologies that allow for sharper approximations of confidence regions.

4. Implementation details and complete algorithm

In this Section, we address important practical aspects needed for the implementation of the MLMC methodology presented in this work and eventually offer pseudo-code of the complete MLMC algorithm. In fact, here we present a unified framework for the estimation of both the expectation $\mathbb{E}[Q]$ and any order central moment $\mu_p(Q)$ of a random variable Q subject to prescribed mean squared error tolerance. As the central moment μ_p is trivially zero for $p = 1$, it will be convenient to denote by $\mathtt{m}_1^{\mathtt{MLMC}}$ the MLMC estimator for $\mathbb{E}[Q]$. Specifically, equation (4) defines the MLMC central moment estimator for any non-trivial order $p > 1$. For $p = 1$ we still use the definition in equation (4) but with a slight abuse of notation by setting $h_1(\boldsymbol{Q}_{N,M}) := \frac{1}{N}\sum_{i=1}^{N} Q_M(\omega_i)$ to denote the sample average operator, so that equation (4) yields the usual MLMC estimator of the expected value for $p = 1$.

In the absence of theoretical estimates for the rates and constants that characterize the bias and statistical error decays as well as the cost model for the problem under investigation (cf. Prop. 3.1), these rates and constants need to be estimated as they are required to optimally tune the MLMC method. That is, to be able to compute the optimal number of levels and sample sizes, a common practice is to perform an initial *screening* procedure. Such a screening procedure consists, for example, of the evaluation of a predefined number of $\overline{N}$ realizations on few (coarse) levels $\{0,\ldots,\overline{L}\}$. Based on these simulations, it is possible to fit these rates and constants (e.g. via a least squares procedure), which then determine the models for the bias, statistical error, and cost per sample.

Once the rates and constants are determined, the pivotal step for achieving the theoretical complexity of the MLMC method subject to a prescribed mean squared error tolerance, is

the choice of both the number of levels L and the sample size N_ℓ required on each level $0 \leq \ell \leq L$. To determine these parameters a precise estimation of the mean squared error (MSE),

$$\mathrm{MSE} = \mathbf{B}^2 + \mathbf{SE} \,,$$

specifically of its two error contributions bias (**B**) and statistical error (**SE**), is required as described in Sect. 3. In order to present a general procedure for the unified MLMC approach to both the expectation and central moments, we recall that

$$\Delta_\ell h_p = h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_\ell}\big) - h_p\big(\boldsymbol{Q}^\ell_{N_\ell,M_{\ell-1}}\big) \,, \quad h_p(\boldsymbol{Q}_{N,M}) = \begin{cases} \frac{1}{N}\sum_{i=1}^N Q_M(\omega_i) \,, & \text{if } p = 1 \,, \\ p\text{-th } h\text{-statistic} \,, & \text{if } p > 1 \,. \end{cases} \tag{9}$$

In practice the bias contribution **B** is thus estimated by

$$\mathbf{B} \approx |\Delta_L h_p| \ . \tag{10}$$

On the other hand, the statistical error **SE** is approximated by

$$\mathbf{SE} \approx \sum_{\ell=0}^{L} \frac{\mathcal{V}_{\ell,p}}{N_\ell} \,, \tag{11}$$

where $\mathcal{V}_{\ell,p}$ denotes the estimated variance $\mathbb{V}ar\big[\Delta_\ell h_p\big]$ on level ℓ. Specifically, we use $\mathcal{V}_{\ell,p} = \hat{V}_{\ell,p}$ on those levels ℓ for which simulations have been run during the screening procedure (i.e. $\ell \leq \overline{L}$). Here, $\hat{V}_{\ell,p}$ is the unbiased sample-based variance estimator introduced in Sect. 3.1. On levels ℓ for which no sample exists yet (i.e. for $\ell > \overline{L}$), we extrapolate the fitted model and use $\mathcal{V}_{\ell,p} = c_\beta {M_\ell}^{-\beta}$ as an estimator. To achieve a prescribed mean squared error of ε^2, we thus require

$$\mathbf{B} \leq \sqrt{1-\theta}\varepsilon \,, \tag{12a}$$

$$\mathbf{SE} \leq \theta\varepsilon^2 \,, \tag{12b}$$

where we have additionally introduced a splitting parameter $\theta \in (0,1)$ to offer the possibility of weighting the two MSE contributions differently. Specifically, the bias constraint (12a) is satisfied for $L \in \mathbb{N}$ such that

$$M_L \geq \left(\frac{\sqrt{1-\theta}\varepsilon}{c_\alpha}\right)^{-\frac{1}{\alpha}} \,, \tag{13}$$

in view of Prop. 3.1(i). Moreover, the theoretical complexity result in Prop. 3.1 also suggests that the statistical error constraint (12b) is satisfied with optimal complexity by selecting the sample size $N_\ell \in \mathbb{N}$ on level ℓ as

$$N_\ell = \left\lceil \frac{1}{\theta\varepsilon^2}\sqrt{\frac{\mathcal{V}_{\ell,p}}{C_\ell}} \sum_{k=0}^{L} \sqrt{C_k \mathcal{V}_{k,p}} \right\rceil \,, \quad \ell = 0, 1, \ldots, L \,, \tag{14}$$

where $C_\ell = \mathrm{cost}(Q_{M_\ell})$.

In Algorithm 1 we provide a detailed pseudo-code of the full MLMC algorithm, which is based on the discussion above. There SOLVE_ℓ denotes a "black-box" solver that, for a

Algorithm 1: MLMC Algorithm for the expectation and central moments of order p.

SCREENING$(\overline{N},\ \overline{L},\ p,\ \varepsilon_r,\ \theta)$
- **for** $\ell = 0 : \overline{L}$ **do**
  - **for** $i = 0 : \overline{N}$ **do**
    - Generate random sample: $\omega_{i,\ell}$
    - $Q_{M_\ell}(\omega_{i,\ell}) \leftarrow \mathsf{SOLVE}_\ell\ (\omega_{i,\ell})$
    - $Q_{M_{\ell-1}}(\omega_{i,\ell}) \leftarrow \mathsf{SOLVE}_{\ell-1}\ (\omega_{i,\ell})$
  - $\Delta_\ell h_p = h_p\big(\boldsymbol{Q}^\ell_{\overline{N},M_\ell}\big) - h_p\big(\boldsymbol{Q}^\ell_{\overline{N},M_{\ell-1}}\big)$, where $h_p(\boldsymbol{Q}_{N,M})$ as in (9)
- **estimate** $\varepsilon = \varepsilon_r \cdot \mathtt{m}_p^{\mathtt{MLMC}}[Q]$
- **estimate** $\{C_\ell\}_{\ell=0}^{\overline{L}},\ \{\mathcal{V}_{\ell,p}\}_{\ell=0}^{\overline{L}}$
- **compute** $\mathcal{P} = \{c_\alpha, c_\beta, c_\gamma, \alpha, \beta, \gamma\}$ using LS fit
- **extrapolate** $\{C_\ell\}_{\ell>\overline{L}},\ \ \{\mathcal{V}_{\ell,p}\}_{\ell>\overline{L}}$
- **compute** L using (13) and N_ℓ using (14)
- **return** $L,\ \{N_\ell\}_{\ell=0}^{L}$

MLMC$(L,\ \{N_\ell\}_{\ell=0}^{L},\ p)$
- **for** $\ell = 0 : L$ **do**
  - **for** $i = 0 : N_\ell$ **do**
    - Generate random sample: $\omega_{i,\ell}$
    - $Q_{M_\ell}(\omega_{i,\ell}) \leftarrow \mathsf{SOLVE}_\ell\ (\omega_{i,\ell})$
    - $Q_{M_{\ell-1}}(\omega_{i,\ell}) \leftarrow \mathsf{SOLVE}_{\ell-1}\ (\omega_{i,\ell})$
  - **compute** $\mathcal{V}_{\ell,p}$
- **compute** $\Delta_L h_p$ as in (9),
- **return** $\mathtt{m}_p^{\mathtt{MLMC}}[Q]$, MSE $= \mathbf{B}^2 + \mathbf{SE}$,

given realization ω_i of the random parameters, returns the approximation $Q_{M_\ell}(\omega_i)$ on the discretization level ℓ. For the sake of completeness, the pseudo-code also contains a possible screening procedure. Eventually, Algorithm 1 returns the MLMC estimator $\mathtt{m}_{\mathtt{p}}^{\mathtt{MLMC}}[Q]$ of the QoI for a prescribed mean squared error tolerance. We emphasize that the implementation presented here takes as an input a *relative* MSE tolerance ε_r, which is related to the commonly used absolute MSE tolerance ε via

$$\varepsilon = \varepsilon_r \begin{cases} \mathbb{E}[Q]\ , & \text{if } p = 1\ , \\ \mu_p(Q)\ , & \text{if } p > 1\ . \end{cases}$$

The unified MLMC framework described in this Section, and much more, is available via our Python library `cmlmc-py`.

5. Numerical Experiments

In this section we apply the introduced multilevel Monte Carlo technique to various examples. In fact, we have used Algorithm 1 with $\theta = 1/2$ in all simulations that follow. We begin by scrutinizing the methodology for rather simple toy problems for which exact (or

	central moment $\mu_p(Q)$		
$\mathbb{E}(Q)$	$p=2$	$p=3$	$p=4$
1.045058356	2.16660855	5.6966642	31.191899

Table 2: Reference values for the expected value $\mathbb{E}[Q]$ and various central moments $\mu_p(Q)$ for the QoI Q derived form the geometric Brownian motion SDE.

highly accurate) solutions are easily available; see Sects. 5.1 and 5.2. Then we move on to study a more challenging problem in Sect. 5.3, namely the one of a transonic airfoil under operational and/or geometric uncertainties.

5.1. Stochastic differential equation model: a financial option

Let us begin with a simple example involving a stochastic differential equation (SDE). Specifically, we consider the case that the SDE models (example borrowed from [10, Sect. 5]) a financial call option with the asset being a geometric Brownian motion, viz.

$$dS = rS\,dt + \sigma S\,dW\ , \quad S(0) = S_0\ . \tag{15}$$

Here, r, σ, and S_0 are given positive numbers. For this asset we are interested in quantifying the uncertainties in the "discounted payoff", so that we set the quantity of interest Q as

$$Q := e^{-rT}\max\big(S(T) - K, 0\big)\ ,$$

where $K > 0$ denotes the agreed strike price and $T > 0$ the pre-defined expiration date. Due to the fact that the solution to (15) at time T, i.e. $S(T)$, is a log-normally distributed random variable with mean $S_0 e^{rT}$ and variance $S_0{}^2 e^{2rT}\big(e^{\sigma^2 T} - 1\big)$, it is straightforward to compute highly accurate approximations to statistics of Q. In fact, Table 2 lists approximated reference values for the expected value and for the first three central moments of Q corresponding to the parameter values $r = \frac{1}{20}$, $\sigma = \frac{1}{5}$, $T = 1$, $K = 10$, and $S_0 = 10$. These reference values were obtained using a high precision numerical quadrature.

For the numerical experiments based on the multilevel Monte Carlo method that will follow, we discretize the SDE (15) via the Milstein method.

$$S_\ell^{n+1} = S_\ell^n + \delta_\ell r S_\ell^n + \sigma S_\ell^n \sqrt{\delta_\ell}\xi_n + \frac{\sigma^2}{2}\delta_\ell S_\ell^n\big(|\xi_n|^2 - 1\big)\ , \quad S_\ell^0 = S_0\ ,$$

so that $S_\ell^{n+1} \approx S(n\delta_\ell)$, where $\delta_\ell = 2^{-\ell}T$ and $(\xi_n)_{n\geq 0}$ denotes a sequence of i.i.d. standard normally distributed random variables. That is, we employ a discretization with a nested grid hierarchy with $M_\ell = T/\delta_\ell = 2^\ell$ DOFs, which corresponds to the number of time steps needed to integrate the SDE from time $t = 0$ to the final time T.

In order to validate the MLMC methodology discussed in this work, we provide in Table 3 a sample based estimation of the MSE of the MLMC estimators, using 100 independent repetitions of the algorithm. Specifically, the table offers a comparison between the required relative root MSE tolerance and the sample based root mean squared error achieved by Algorithm 1 for both the expected value and the first three central moments for various

Tol	$\mathtt{m}_1^{\mathtt{MLMC}}$	$\mathtt{m}_2^{\mathtt{MLMC}}$	$\mathtt{m}_3^{\mathtt{MLMC}}$	$\mathtt{m}_4^{\mathtt{MLMC}}$
$\varepsilon_r = 0.1$	0.0647	0.0857	0.0785	0.0991
$\varepsilon_r = 0.05$	0.0435	0.0429	0.0415	0.0495
$\varepsilon_r = 0.025$	0.0237	0.0217	0.0231	0.0223
$\varepsilon_r = 0.01$	0.0087	0.0099	0.0091	0.0083

Table 3: Sample estimate of relative root MSE based on 100 repetitions of the MLMC algorithm for computing $\mathtt{m}_p^{\mathtt{MLMC}}$ for different relative tolerance requirements.

tolerances. The results in Table 3 demonstrate that the MLMC implementation described in the previous Section does indeed provide estimators that satisfy the tolerance requirement. Additionally, in the top row of Figure 1 we show the actual computed values of these 100 repetitions of the MLMC algorithm (red circles) compared with the reference solution (green stars). To quantify the range of the MLMC estimators, we also indicate the 90% confidence intervals based on a Chebyshev bound (blue bars; see Eq. (8)) in these plots.

The bottom row in Figure 1 presents the corresponding MLMC hierarchies (both number of levels and sample size per level) required to achieve prescribed relative tolerance requirements when estimating the expectation and various central moments of the QoI Q. It is interesting to observe that the computational cost required to compute central moments is proportional to that for the expectation up to a multiplicative constant. The latter can be further observed in Figure 2, where we plot the computed bias and variance of the estimators, respectively, for various tolerance demands. It can be inferred that the decay rate for the estimator's bias and variances is the same, while the constants are increasing with increasing p.

5.2. Elliptic PDE in two spatial dimensions

We consider a random Poisson equation in two spatial dimensions,

$$-\Delta u = f \ , \quad \text{in } D = (0,1)^2 \ , \tag{16}$$

with homogeneous Dirichlet boundary conditions. Here, the forcing term f is given by

$$f(x) = -K\xi({x_1}^2 + {x_2}^2 - x_1 - x_2) \ ,$$

with ξ being a non-negative random variable and $K > 0$ a positive constant. For this forcing term the solution to the PDE can be computed explicitly and reads $u(x_1, x_2) = K\xi x_1 x_2(1-x_1)(1-x_2)/2$. As quantity of interest we consider the spatial average of the solution, that is

$$Q := \int_D u \, dx = \frac{K}{72}\xi \ .$$

This explicit representation of Q in terms of the random input ξ to the PDE model (16) allows us to easily compute the exact mean as well as central moments of Q, which we will use to verify the numerical experiments that follow. Specifically, here we use $\xi \sim \text{Beta}(2,6)$ and $K = 432$. Table 4 then lists approximations to the corresponding mean and the first three central moments of Q. For the numerical experiments based on multilevel Monte Carlo method we discretize the PDE (16) using a second order finite difference scheme on a regular

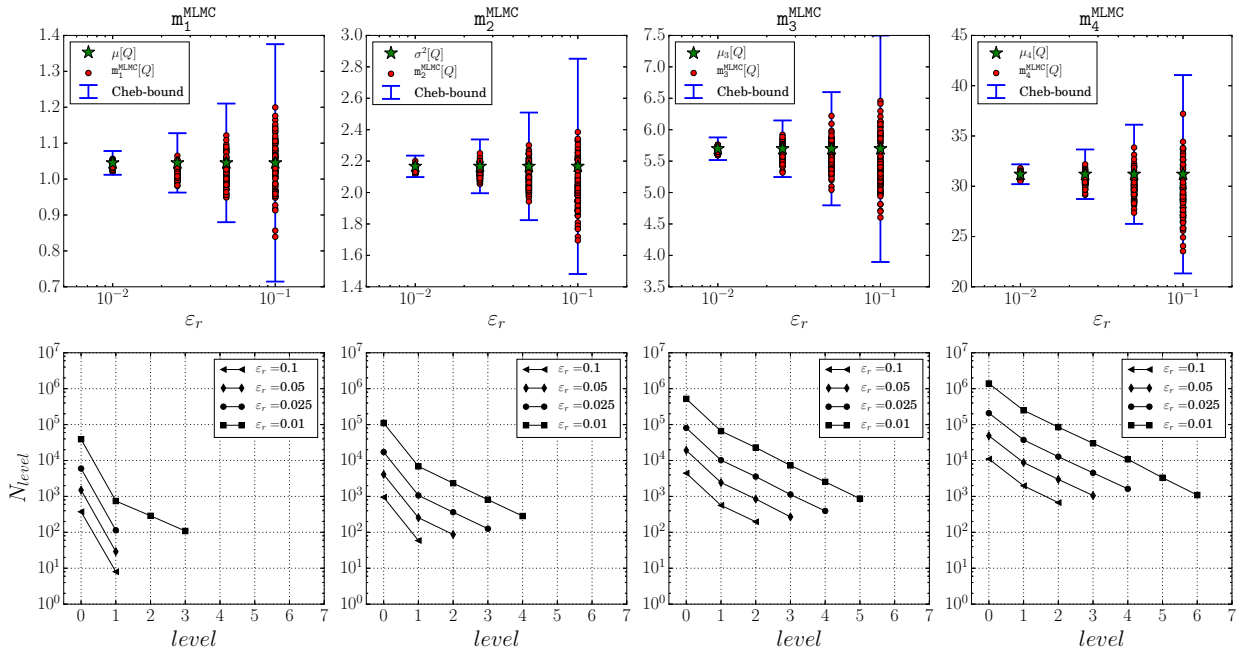


Figure 1: Computed values of 100 repetitions of the MLMC algorithm compared with the reference solution (first row) and MLMC hierarchies (number of levels and sample size per level) required to achieve prescribed relative tolerance requirements when estimating the expectation and various central moments of the QoI Q (second row) for the SDE problem.

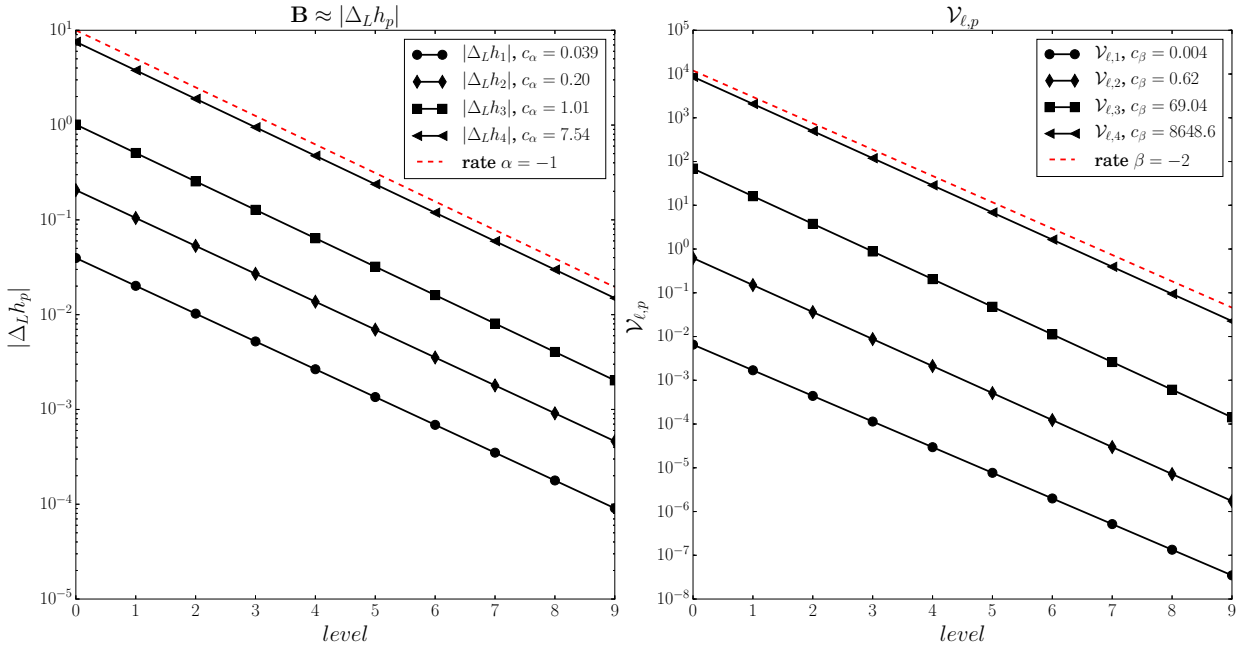


Figure 2: Decay rates for the bias and variances of MLMC estimator with increasing p for the SDE problem.

grid. That is, we employ a nested grid hierarchy with $M_\ell = (5 \cdot 2^\ell - 2)^2$ DOFs, which correspond to the values of the solution u at grid points that are not on the boundary ∂D.

	central moment $\mu_p(Q)$		
$\mathbb{E}(Q)$	$p=2$	$p=3$	$p=4$
1.5	0.75	0.45	1.748863636

Table 4: Reference values for the expected value $\mathbb{E}[Q]$ and the first three central moments $\mu_p(Q)$ for the QoI Q derived form the random Poisson problem.

As for the previous example, in Table 5 we present the (sample based) root mean squared errors obtained by repeating the MLMC algorithm for the expectation and central moments 100 times and for various tolerances. Also for this example we find that the MLMC imple-

Tol	$\mathtt{m}_1^{\mathtt{MLMC}}$	$\mathtt{m}_2^{\mathtt{MLMC}}$	$\mathtt{m}_3^{\mathtt{MLMC}}$	$\mathtt{m}_4^{\mathtt{MLMC}}$
$\varepsilon_r = 0.1$	0.0674	0.0616	0.0587	0.0777
$\varepsilon_r = 0.05$	0.0350	0.0401	0.0351	0.0259
$\varepsilon_r = 0.025$	0.0182	0.0183	0.0156	0.0206
$\varepsilon_r = 0.01$	0.0069	0.0078	0.0062	0.0084

Table 5: Sample estimate of relative root MSE of 100 repetitions of the MLMC estimators $\mathtt{m}_p^{\mathtt{MLMC}}$ for different relative tolerance requirements.

mentation does indeed satisfy the required tolerance goals.

In the top row of Figure 3 the actual computed values for 100 repetitions of the MLMC algorithm (red circles) are compared with the reference values (green stars). Also for this example we observe an accurate estimation within the imposed tolerance goal and within the confidence region (blue bars, 90% confidence; see Eq. (8)). In the second row of Figure 3 we report the hierarchies required to achieve the prescribed tolerances. As it is possible to observe the number of levels and samples per level (and hence the cost) increase consistently with the central moment we are computing. Such can be also inferred by looking at the decays of the bias and variance of the MLMC estimators for moments presented in Figure 4. These plots moreover confirm the observation from the previous example, namely that the decay rates for the estimator's bias and variance are the same for different values of p, and only the constants vary.

5.3. Transonic Airfoil: 2d

We consider hereafter a transonic supercritical RAE-2822 airfoil [11, 20], which has become a standard test-case for transonic flows, subject to both operating and geometric uncertainties.

The fluid flowing around an airfoil generates a local force on each point of the body. The normal and tangential components of such force are the pressure and the shear stress. By integrating the force and stress distribution around the surface of the airfoil we obtain a total force F and a moment M about a reference point (so called center of pressure). The parallel and perpendicular component of F with respect to the free-stream direction M_∞ are the lift L and drag D forces respectively. Figure 5(a) shows a sketch of this concept. For an airfoil shape with surface S we define the following lift, drag, and moment dimensionless

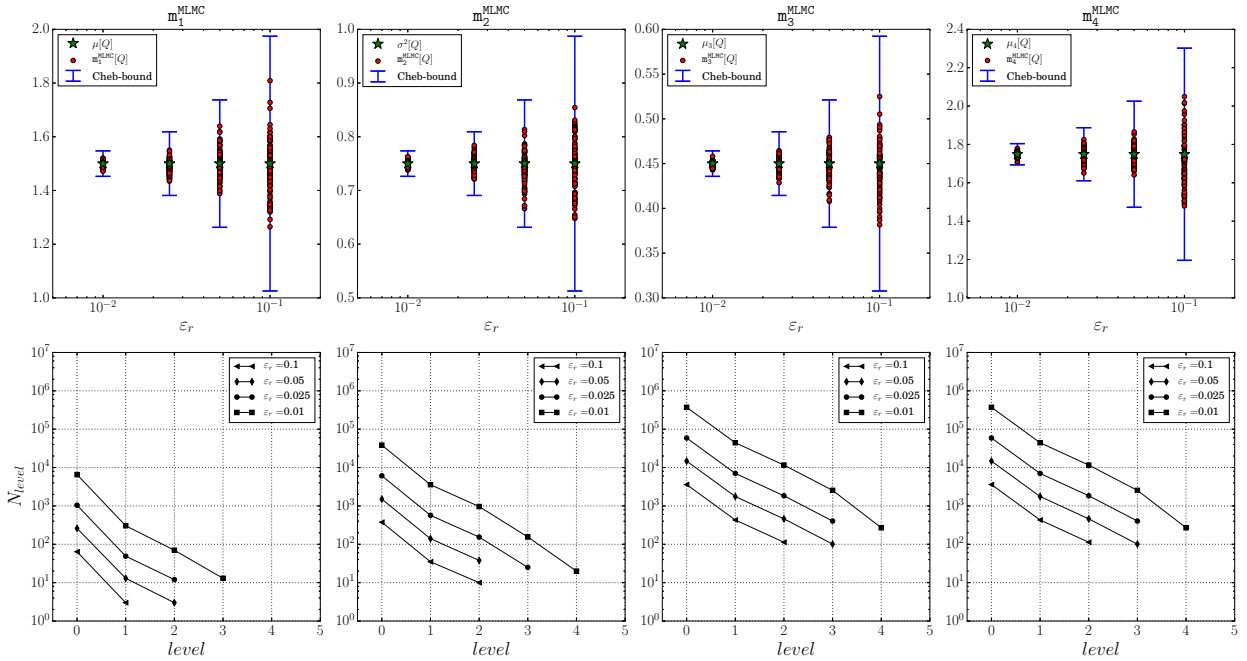


Figure 3: Computed values of 100 repetitions of the MLMC algorithm compared with the reference solution (first row) and MLMC hierarchies (number of levels and sample size per level) required to achieve prescribed relative tolerance requirements when estimating the expectation and various central moments of the QoI Q (second row) for the Elliptic PDE problem.

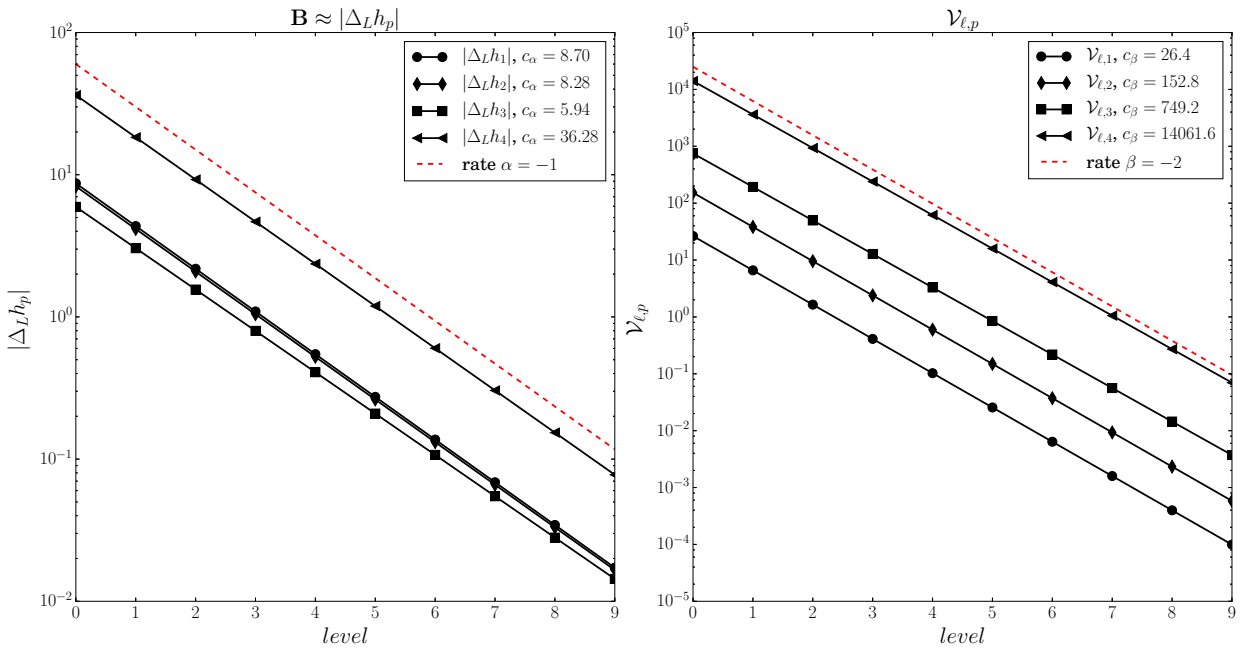
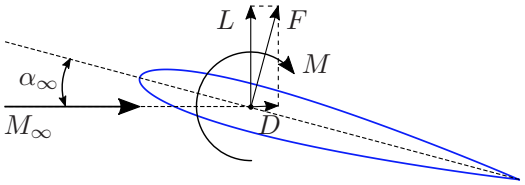


Figure 4: Decay rates for the bias and variances of MLMC estimator with increasing p for the Elliptic PDE problem.

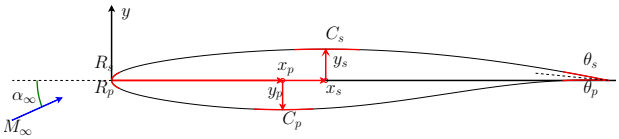
coefficients:

$$C_L = \frac{L}{q_\infty S}\,, \qquad C_D = \frac{D}{q_\infty S}\,, \qquad \text{and} \qquad C_M = \frac{M}{q_\infty S L_{ref}}\,, \tag{17}$$

respectively. Here, $q_\infty = \frac{1}{2} M_\infty^2 \gamma_g p_\infty$ denotes the dynamic pressure and $\gamma_g = 1.4$ is the ratio of specific heats of the gas. As we are considering 2D normalized airfoils we set the reference length $L_{ref} = 1$ and the reference surface $S = 1$.



(a) Aerodynamic forces and moments acting on the airfoil.



(b) Geometry of the RAE-2822 transonic airfoil and PARSEC parameters that define the geometry of the airfoil.

Figure 5: Description of the airfoil set-up considered in this work.

The nominal geometry of the RAE-2822 airfoil is defined by a set of PARSEC parameters (see [18] for details). The advantage of the PARSEC approach over other parametrizations (i.e. Bezier, NURBS, FFD) is that we can easily perturb the geometrical parameters on the suction and pressure side of the airfoil (see Figure 5(b)) which are most relevant for the study that follows. Among other things, Table 6 summarizes the geometric definition of the airfoil as well as the set of operating parameters for three different flow conditions considered here. Specifically, CASE-6 denotes the mild transonic case (corresponding to experimental case 6 from AGARD [20]), CASE-S is a subsonic case with $M_\infty = 0.6$, and CASE-R is a higher Reynolds number case.

	Name	Nominal value			Uncertainty
		CASE-6	**CASE-S**	**CASE-R**	
Operating	Re_c	$6.5e6$	$6.5e6$	$10e6$	–
	M_∞	0.729	0.6	0.729	$\mathcal{B}(4, 2, 0.05, M_\infty - 0.037)$
	α_∞	2.31°	2.31°	2.31°	$\mathcal{B}(4, 2, 0.2, 2.16)$
Geometric	R_p	$8.60311920e-03$			$\mathcal{U}(98\%, 102\%)$
	R_s	$8.36101985e-03$			$\mathcal{U}(98\%, 102\%)$
	x_p	$3.44224863e-01$			$\mathcal{U}(98\%, 102\%)$
	x_s	$4.31244633e-01$			$\mathcal{U}(98\%, 102\%)$
	y_p	$-5.88259641e-02$			$\mathcal{U}(98\%, 102\%)$
	y_s	$6.30175650e-02$			$\mathcal{U}(98\%, 102\%)$
	C_p	$7.03608884e-01$			$\mathcal{U}(98\%, 102\%)$
	C_s	$-4.30110180e-01$			$\mathcal{U}(98\%, 102\%)$
	θ_p	$-2.06545825e+00$			$\mathcal{U}(98\%, 102\%)$
	θ_s	$-1.15335351e+01$			$\mathcal{U}(98\%, 102\%)$

Table 6: Operational and geometrical parameters as well as a description of the uncertainties for the RAE-2822 airfoil.

In what follows, we consider the RAE-2822 airfoil in three different operating regimes with increasing number of uncertain parameters. Specifically, we use the letter G to denote

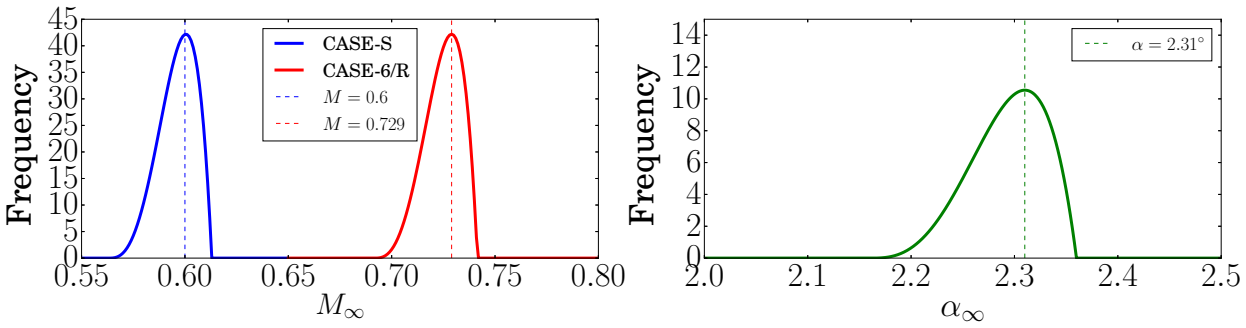


Figure 6: Exemplary probability density functions of two uncertain operating input parameters for the random RAE-2822.

stochastic simulations where we consider only *geometric* uncertainties (i.e. 10 random input parameters), O to denote the regime with only *operating* uncertainties (i.e. two random input parameters, namely the angle of attack α_∞ and the Mach number M_∞, see Figure 6) and OG to denote the setting with *geometric plus operating* uncertainties (i.e. 12 uncertain input parameters). All uncertainties and reference nominal operating and geometric parameters are presented in Table 6. The operating uncertainties are modeled as beta distributions denoted by $\mathcal{B}(a, b, s, loc)$, where a and b are the distribution parameters. As the beta distribution is defined on the $[0, 1]$ interval, the parameters s and loc are used to scale and shift the distribution's support, respectively. On the other hand the geometric uncertainties are modeled as uniform distributions, denoted by $\mathcal{U}(x_{low}, x_{up})$ with $x_{low} < x_{up}$ denoting range of the support. In Table 6 x_{low} and x_{up} are given as percentages of the nominal value. The types and ranges of uncertainties for this model problem are representative of a flight condition with natural atmospheric gusts that affect both the angle of attack and the Mach number. Additionally, the geometrical uncertainties are reasonably accounting for manufacturing tolerances and shape deformation of a airfoil due to different loadings on an aircraft wing (aeroelastic twist).

For the numerical study that follows we use the MSES collection of programs for the analysis of airfoils (see [7] for detail) as deterministic 'black-box' solver. The MSES collection solves the steady Euler equations with a finite volume discretization over a streamline grid and is coupled, via the displacement thickness, with a two-equation integral solver for the viscous regions of the boundary layer and trailing wakes. The performance of this 'black-box' solver, when using a 5-levels structured MLMC grid hierarchy, is summarized in Table 7. Specifically, there the features of the grid levels, along with the average computational time $CTime$ required to compute one deterministic simulation (on one CPU) are shown.

Based on the problem description of the uncertain airfoil problem considered here, in the following study we apply the developed MLMC estimator for central moments to various aerodynamic performance parameters. In order to present the estimated expectations and central moments estimators for the three different cases (CASE-6, CASE-S, and CASE-R) with increasing number of uncertain parameters (G, O, and OG) in a compact and informative way, we introduce in Figure 7 a set of bars that are designed to provide the relevant information. Specifically, there, the mean, the standard deviation, the skewness, and the kurtosis of different QoIs related to the airfoil, such as lift coefficient C_L, drag coefficient

Level	Airfoil nodes	Cells	$CTime[s]$
$L0$	47	1739	1.9
$L1$	71	2627	3.2
$L2$	107	3959	5.7
$L3$	161	5957	7.5
$L4$	243	8991	14.7
$L5$	365	13505	17.9

Table 7: MLMC 5-levels grid hierarchy for the RAE2822 problem.

C_D, moment coefficient C_M, and lift-drag ratio L/D, are presented. Moreover, we compare the deterministic value (obtained with nominal geometric and operating parameters) of a QoI (dashed black lines) with a classical mean plus/minus two standard deviation interval (black bars). The red bars identify the skewness corrected mean plus/minus two standard deviation, where the skewness correct mean is given by $\mu + \gamma$. Moreover, the triangles define the kurtosis: yellow inward triangles identify the *platykurtic* distributions while red outward triangles denote *leptokurtic* ones. A distribution is called platykurtic, if the kurtosis $Kurt < 3$, which means that the distribution has thinner tails than a Gaussian distribution. Similarly, a distribution is called leptokurtic if $Kurt > 3$, which implies fatter tails.

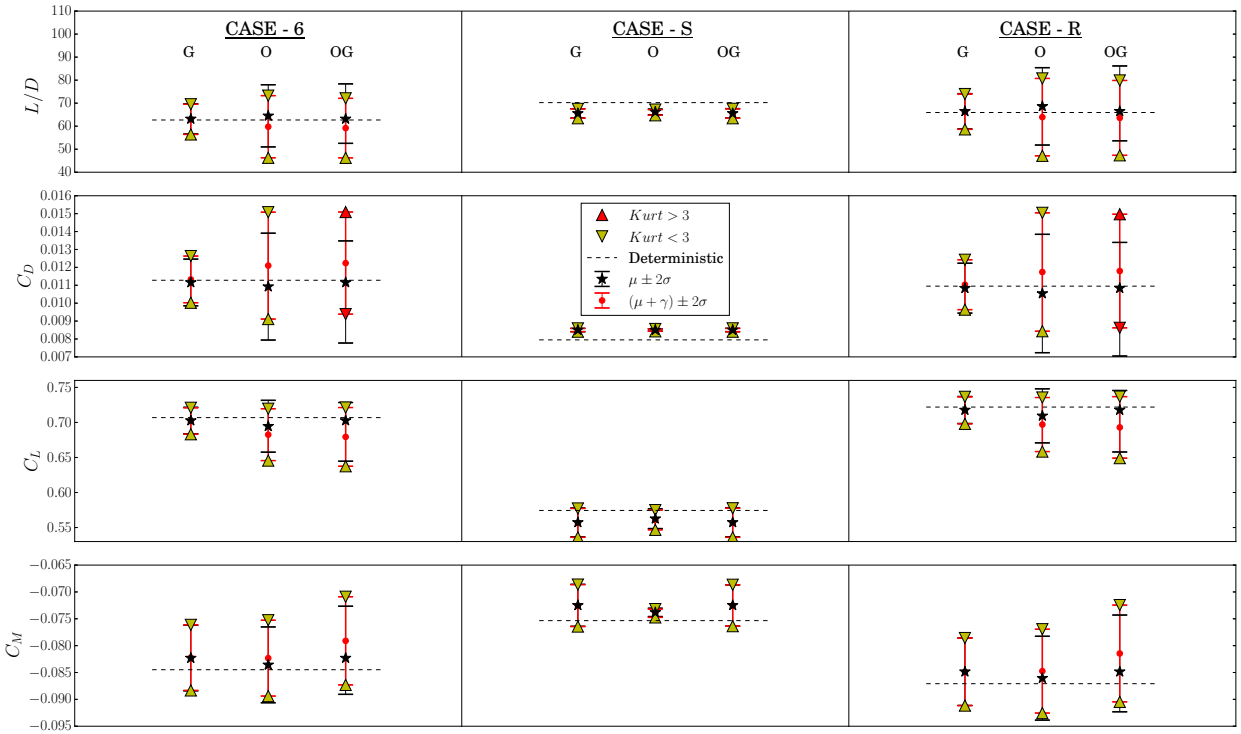


Figure 7: Estimated expectations and central moments for the lift-drag ratio L/D, drag coefficient C_D, lift coefficient C_L and moment coefficient C_M of the RAE-2822 airfoil in different operating conditions and considering different uncertainties scenarios.

It is interesting to observe in Figure 7 the effects of uncertainties on the performance

parameters in the three different flow conditions. The two transonic cases CASE-6 and CASE-R are the most sensitive to uncertainties due to the appearance of shock waves in different regions on the airfoil upper side (see also Figure 8). In such cases we additionally notice that the drag coefficient C_D becomes leptokurtic in the presence of both operating and geometric uncertainties, indicating that the distribution is heavy-tailed. We believe that this is due to the appearance of separation bubbles in the front part of the airfoil and stronger shock waves, but further investigations are necessary to confirm this hypothesis. The variability in the forward part of the airfoil can be observed also in the green C_p plot in Figure 8. Further investigation are needed to confirm this hypothesis.

Additionally, in Figure 8 we compare the pressure coefficients C_p of the RAE 2822 airfoil in the different conditions and uncertainty scenarios introduced above as well as the reconstructed lift-drag ratio L/D distributions computed from the statistical moments using the Gram–Charlier series of type A PDF approximation [22]. The latter is a formal series expansions in terms of a known distribution, most commonly with respect to a Normal distribution. Using this approach, an unknown density f can be approximated by

$$\hat{f}(x) := \frac{1}{\sqrt{2\pi\mu_2}} \exp\left(-\frac{(x-\mu)^2}{2\mu_2}\right)\left(1 + \frac{\mu_3}{3!{\mu_2}^{3/2}} H_3\left(\frac{x-\mu}{\sqrt{\mu_2}}\right) + \frac{\mu_4 - 3{\mu_2}^2}{4!{\mu_2}^2} H_4\left(\frac{x-\mu}{\sqrt{\mu_2}}\right)\right),$$

where $H_3(x) = x^3 - 3x$ and $H_4(x) = x^4 - 6x^2 + 3$ are Hermite polynomials. Although $\hat{f}$ may formally not be a proper density as it is not guaranteed to be positive, it nonetheless offers an easy to compute density approximation, based on the MLMC estimators for μ, μ_2, μ_3, and μ_4. By looking at the C_p profiles and the reconstructed PDF approximations we can further observe the sensitivity of the airfoil on operating and geometric parameters in the three different flow cases as previously noticed in Figure 7. It is worth underlining that the PDF presented here are simply reconstructed from the first four central moments computed with the MLMC method. The background histograms are obtained from a MC simulation with 1000 samples on the finest level. A more efficient and accurate procedure to compute directly the PDF of QoIs will be presented in the forthcoming Part II of this work.

Finally, in Figure 9 we present the computational complexity in CPU hours required to achieve a certain tolerance requirement. Specifically, the complexities of the MLMC method and the classic MC approach are compared for approximating the expectation and central moments for the L/D QoI in CASE-6 with both operating and geometric uncertainties. The dashed lines indicate the computational complexity predicted by the theory. We immediately observe a significant speedup of the MLMC method compared to the MC method. Practically speaking, by employing a cluster node with 28 CPUs we are able to compute the first four central moments of the airfoil problem and guarantee a relative tolerance of 1% (i.e. $\varepsilon_r = 0.01$) in 3.6 $[h]$ with our MLMC implementation, while we would need to invest 14.8 $[days]$ with the classic MC method.

6. Conclusion

In this work we have introduced an extension to the multilevel Monte Carlo (MLMC) method that allows to efficiently compute central statistical moments of a random system output's quantity of interest. The key feature of our procedure is the use of h-statistics as

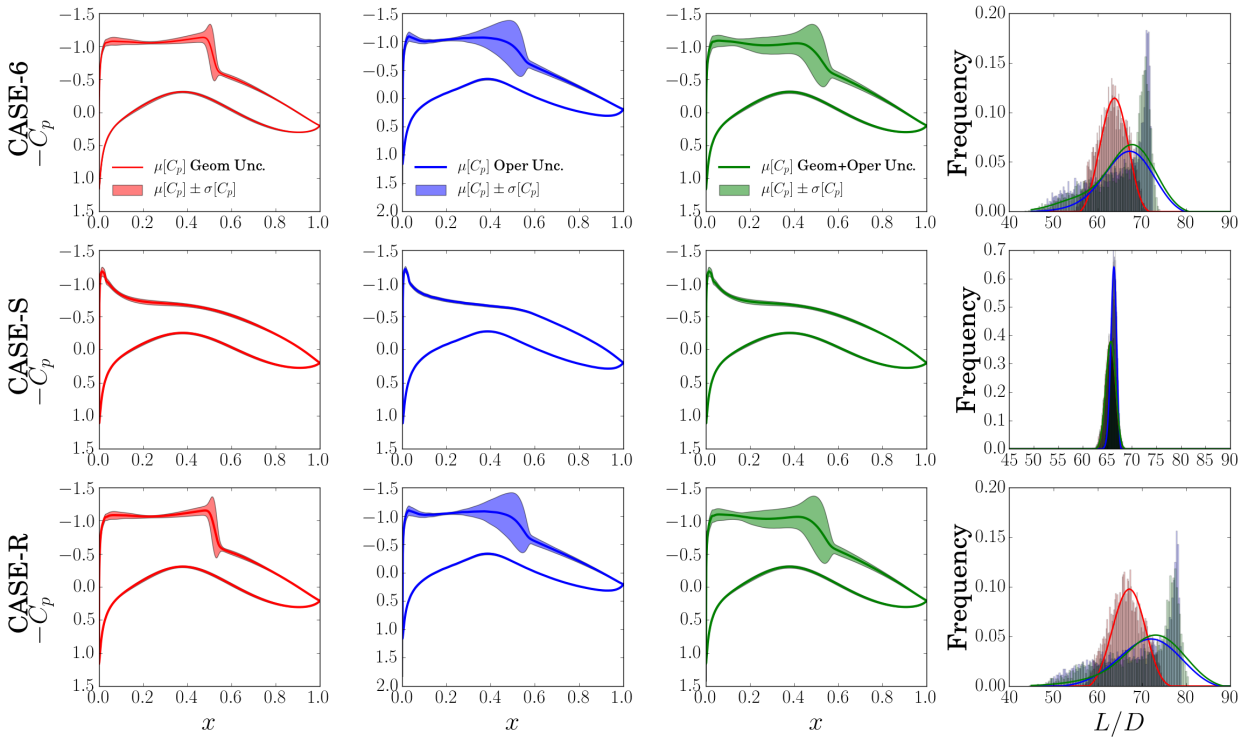


Figure 8: Pressure coefficients C_p of the RAE 2822 airfoil in the different conditions and uncertainty scenarios and the reconstructed lift-drag ratio L/D distributions.

unbiased central moment estimators with minimal variance for the level-wise contributions. Using an extensive set of numerical benchmark examples, we have demonstrated that the proposed MLMC estimator based on h-statistics satisfy the mean squared error tolerance requirement and requires a computational cost proportional to that for the MLMC estimation of an expectation (up to a multiplicative constant). Additionally we observed that the decay rate for the estimator's bias and variances is the same for arbitrary order central moment μ_p, while the constants are increasing with increasing p.

For the numerical experiments related to the airfoil problem, we observed that central moments can also provide relevant information regarding random variable distribution in view of decision making processes and optimization under uncertainty approaches. In fact, we tested a distribution reconstruction approach bases on series expansion form the statistical moment (Gram–Charlier approximation) and observed that the approximation of the distribution is not always satisfactory. The reconstruction is not guaranteed to be a proper probability distribution and seems sometimes to lead to large inaccuracies in capturing asymmetric behaviors and heavy tails. In order to overcome these issues, the forthcoming Part II of this work will address another extension of the general MLMC approach to accurately approximate distributions and risk measures.

Acknowledgments

M. P. and F. N. acknowledge funding from the EU-FP7 framework program project UMRIDA under grant agreement no. ACP3-GA-2013-605036. This work has also been supported

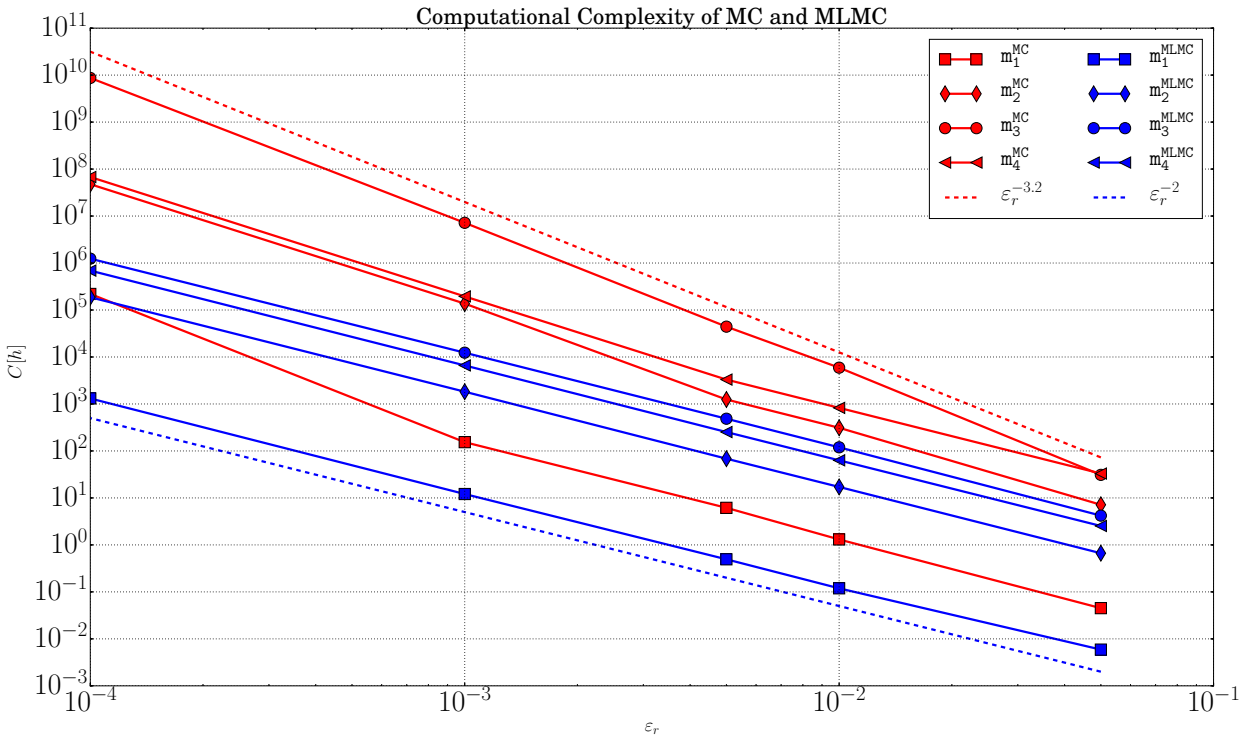


Figure 9: Computational complexity of MC and MLMC (in CPU hours) required to achieve a certain tolerance requirement for the first central moments.

by the Center for Advanced Modeling Science (CADMOS).

AppendixA. Unbiased variance estimator for $p = 4$ for MC method

Below we report the closed-form expressions of the unbiased estimators $\hat{V}_4/N$ for $\mathbb{V}\mathrm{ar}\big[h_4(\boldsymbol{Q}_{N,M})\big] = V_4/N$, $S_a \equiv S_a(\boldsymbol{Q}_{N,M})$:

$$
\begin{aligned}
\frac{\hat{V}_4}{N} = \frac{1}{(N-7)(N-6)(N-5)(N-4)(N-3)^2(N-2)^2(N-1)^2N^2}\{&-72(2N^3-21N^2+79N-105)S_1^8 + 288N(2N^3-21N^2+79N-105)S_2S_1^6 \\
&- 48(7N^5-85N^4+443N^3-1199N^2+1734N-1260)S_3S_1^5 - 12(3(17N^5-167N^4+505N^3-61N^2-1734N+1260)S_2^2 \\
&+ N(-13N^5+157N^4-851N^3+2387N^2-3336N+2016)S_4)S_1^4 - 24N(2(-11N^5+134N^4-682N^3+1804N^2-2667N+2142)S_2S_3 \\
&+ (3N^6-40N^5+250N^4-812N^3+1319N^2-972N+252)S_5)S_1^3 + 4(18N(2N^5-11N^4-59N^3+581N^2-1467N+1134)S_2^3 \\
&- 3N(11N^6-114N^5+452N^4-714N^3+77N^2+828N-1260)S_4S_2 - 4(7N^7-114N^6+847N^5-3603N^4+9532N^3-15867N^2+15318N-7560)S_3^2 \\
&+ N(7N^7-106N^6+778N^5-3136N^4+7387N^3-10798N^2+9396N-3528)S_6)S_1^2 \\
&- 8(N(N^6-15N^5+109N^4-417N^3+934N^2-1332N+936)S_7(N-1)^2 + 6(2N^7-15N^6-25N^5+627N^4-2785N^3+6120N^2-7344N+3780)S_2^2S_3 \\
&+ N(-6N^7+101N^6-792N^5+3626N^4-10470N^3+18917N^2-19260N+9324)S_3S_4 \\
&- 3N(N^7-7N^6-14N^5+350N^4-1715N^3+3941N^2-4320N+1764)S_2S_5)S_1 - 36(2N^6-17N^5-23N^4+659N^3-2493N^2+3672N-1890)S_2^4 \\
&+ 12N(11N^6-135N^5+581N^4-909N^3+20N^2+540N+252)S_2^2S_4 + 4N(N^3-6N^2+11N-6)S_2(4(N^4-7N^3+2N^2+118N-294)S_3^2 \\
&+ (-13N^4+124N^3-527N^2+1148N-1092)S_6) + N((-N^8+N^7+148N^6-1694N^5+9715N^4-33983N^3+70850N^2-76356N+32760)S_4^2 \\
&+ (N^3-6N^2+11N-6)(N(N^5-11N^4+63N^3-145N^2+248N-156)S_8 - 8(N^5-14N^4+93N^3-346N^2+854N-1092)S_3S_5))\}
\end{aligned}
$$

AppendixB. Unbiased variance estimator on level ℓ for $p = 3$

For the sake of completeness, we present here the closed-form expression for the unbiased estimator of $\mathbb{V}\mathrm{ar}\big(\Delta_\ell h_p\big) \equiv V_{\ell,p}/N_\ell$ for $p = 3$.

$$
\begin{aligned}
\frac{\hat{V}_{\ell,3}}{N_\ell} = {} & \frac{1}{16(N_\ell-5)(N_\ell-4)(N_\ell-3)(N_\ell-2)^2(N_\ell-1)^2{N_\ell}^2}\{-12(3(N_\ell-5)N_\ell+20)S_{0,1}^6 \\
& + 36(2N_\ell S_{2,0}(N_\ell-3)^2 + (3(N_\ell-5)N_\ell+20)(N_\ell S_{0,2}-2S_{1,0}^2))S_{0,1}^4 \\
& + 24N_\ell(-(N_\ell(2N_\ell-9)+11)(N_\ell S_{0,3}-6S_{1,0}S_{1,1}) - 3((N_\ell-3)(N_\ell-2)N_\ell+2)S_{2,1})S_{0,1}^3 \\
& + 3(-6{N_\ell}^2(N_\ell(4N_\ell-21)+29)S_{0,2}^2 + 12N_\ell((8{N_\ell}^2-42N_\ell+58)S_{1,0}^2 \\
& + (2-(N_\ell-3)N_\ell(3N_\ell-10))S_{2,0})S_{0,2} + (N_\ell-2)(N_\ell-1)N_\ell(N_\ell(7N_\ell-15)+20)S_{0,4} \\
& + 3(-12(3(N_\ell-5)N_\ell+20)S_{1,0}^4 + 8N_\ell(N_\ell(5N_\ell-24)+31)S_{2,0}S_{1,0}^2 \\
& - 8(N_\ell-2)(N_\ell-1)^2N_\ell(2S_{1,2}+S_{3,0})S_{1,0} + N_\ell(-8(N_\ell-2)(N_\ell-1)^2S_{1,1}^2 \\
& - 2((N_\ell-1)N_\ell(2N_\ell-7)+12)S_{2,0}^2 + (N_\ell-2)(N_\ell-1)(2(N_\ell(3N_\ell-7)+8)S_{2,2} \\
& + ((N_\ell-1)N_\ell+4)S_{4,0}))))S_{0,1}^2 - 6N_\ell(4S_{2,3}{N_\ell}^5 - 6S_{1,1}S_{1,2}{N_\ell}^4 - 10S_{1,0}S_{1,3}{N_\ell}^4 \\
& - 5S_{0,3}S_{2,0}{N_\ell}^4 - 9S_{2,0}S_{2,1}{N_\ell}^4 - 20S_{2,3}{N_\ell}^4 - 6S_{1,1}S_{3,0}{N_\ell}^4 - 18S_{1,0}S_{3,1}{N_\ell}^4 + 12S_{0,3}S_{1,0}^2{N_\ell}^3 \\
& + 48S_{1,0}S_{1,3}{N_\ell}^3 + 30S_{0,3}S_{2,0}{N_\ell}^3 + 60S_{1,0}S_{1,1}S_{2,0}{N_\ell}^3 + 48S_{1,0}^2S_{2,1}{N_\ell}^3 + 30S_{2,0}S_{2,1}{N_\ell}^3 \\
& + 52S_{2,3}{N_\ell}^3 + 24S_{1,1}S_{3,0}{N_\ell}^3 + 96S_{1,0}S_{3,1}{N_\ell}^3 - 60S_{0,3}S_{1,0}^2{N_\ell}^2 - 96S_{1,0}^3S_{1,1}{N_\ell}^2 \\
& + 90S_{1,1}S_{1,2}{N_\ell}^2 - 106S_{1,0}S_{1,3}{N_\ell}^2 - 35S_{0,3}S_{2,0}{N_\ell}^2 - 324S_{1,0}S_{1,1}S_{2,0}{N_\ell}^2 - 228S_{1,0}^2S_{2,1}{N_\ell}^2 \\
& + 45S_{2,0}S_{2,1}{N_\ell}^2 - 92S_{2,3}{N_\ell}^2 - 6S_{1,1}S_{3,0}{N_\ell}^2 - 210S_{1,0}S_{3,1}{N_\ell}^2 + 72S_{0,3}S_{1,0}^2N_\ell + 504S_{1,0}^3S_{1,1}N_\ell \\
& - 180S_{1,1}S_{1,2}N_\ell + 132S_{1,0}S_{1,3}N_\ell - 30S_{0,3}S_{2,0}N_\ell + 480S_{1,0}S_{1,1}S_{2,0}N_\ell + 276S_{1,0}^2S_{2,1}N_\ell \\
& - 210S_{2,0}S_{2,1}N_\ell + 88S_{2,3}N_\ell - 60S_{1,1}S_{3,0}N_\ell + 228S_{1,0}S_{3,1}N_\ell + 24S_{0,3}S_{1,0}^2 \\
& + (N_\ell-2)(N_\ell-1)^2((N_\ell-1)N_\ell+4)S_{0,5} - 696S_{1,0}^3S_{1,1} + 96S_{1,1}S_{1,2} - 64S_{1,0}S_{1,3} \\
& + 16S_{0,3}S_{2,0} - 72S_{1,0}S_{1,1}S_{2,0} + 48S_{1,0}^2S_{2,1} + 72S_{2,0}S_{2,1} \\
& + S_{0,2}((N_\ell(N_\ell(N_\ell(18-5N_\ell)+13)-90)+40)S_{0,3} + 36((N_\ell-3)(N_\ell-2)N_\ell+2)S_{1,0}S_{1,1} \\
& - 3(N_\ell(3{N_\ell}^3 - 14{N_\ell}^2 + N_\ell + 50) - 16)S_{2,1}) - 32S_{2,3} + 48S_{1,1}S_{3,0} - 96S_{1,0}S_{3,1} \\
& + 3(N_\ell-2)(N_\ell-1)^2((N_\ell-1)N_\ell+4)S_{4,1})S_{0,1} + N_\ell(S_{0,6}{N_\ell}^6 + 6S_{2,4}{N_\ell}^6 - 9S_{2,1}^2{N_\ell}^5 \\
& - 5S_{0,6}{N_\ell}^5 - 12S_{1,1}S_{1,3}{N_\ell}^5 - 12S_{1,0}S_{1,4}{N_\ell}^5 - 6S_{0,4}S_{2,0}{N_\ell}^5 - 18S_{2,0}S_{2,2}{N_\ell}^5 - 30S_{2,4}{N_\ell}^5 \\
& - 36S_{1,1}S_{3,1}{N_\ell}^5 - 36S_{1,0}S_{3,2}{N_\ell}^5 + 12S_{0,4}S_{1,0}^2{N_\ell}^4 - 36S_{1,2}^2{N_\ell}^4 + 13S_{0,6}{N_\ell}^4 + 48S_{1,1}S_{1,3}{N_\ell}^4 \\
& + 60S_{1,0}S_{1,4}{N_\ell}^4 + 72S_{1,1}^2S_{2,0}{N_\ell}^4 + 42S_{0,4}S_{2,0}{N_\ell}^4 + 36S_{1,0}S_{1,2}S_{2,0}{N_\ell}^4 + 144S_{1,0}S_{1,1}S_{2,1}{N_\ell}^4 \\
& + 72S_{1,0}^2S_{2,2}{N_\ell}^4 + 90S_{2,0}S_{2,2}{N_\ell}^4 + 78S_{2,4}{N_\ell}^4 - 36S_{1,2}S_{3,0}{N_\ell}^4 + 216S_{1,1}S_{3,1}{N_\ell}^4 \\
& + 180S_{1,0}S_{3,2}{N_\ell}^4 - 72S_{0,4}S_{1,0}^2{N_\ell}^3 - 216S_{1,0}^2S_{1,1}^2{N_\ell}^3 + 144S_{1,2}^2{N_\ell}^3 + 225S_{2,1}^2{N_\ell}^3 - 23S_{0,6}{N_\ell}^3 \\
& - 72S_{1,0}^3S_{1,2}{N_\ell}^3 - 12S_{1,1}S_{1,3}{N_\ell}^3 - 156S_{1,0}S_{1,4}{N_\ell}^3 - 576S_{1,1}^2S_{2,0}{N_\ell}^3 - 78S_{0,4}S_{2,0}{N_\ell}^3 \\
& - 72S_{1,0}S_{1,2}S_{2,0}{N_\ell}^3 - 648S_{1,0}S_{1,1}S_{2,1}{N_\ell}^3 - 360S_{1,0}^2S_{2,2}{N_\ell}^3 - 90S_{2,0}S_{2,2}{N_\ell}^3 - 138S_{2,4}{N_\ell}^3 \\
& + 144S_{1,2}S_{3,0}{N_\ell}^3 - 324S_{1,1}S_{3,1}{N_\ell}^3 - 468S_{1,0}S_{3,2}{N_\ell}^3 + 156S_{0,4}S_{1,0}^2{N_\ell}^2 + 1224S_{1,0}^2S_{1,1}^2{N_\ell}^2 \\
& - 36S_{1,2}^2{N_\ell}^2 - 324S_{2,1}^2{N_\ell}^2 + 22S_{0,6}{N_\ell}^2 + 288S_{1,0}^3S_{1,2}{N_\ell}^2 - 312S_{1,1}S_{1,3}{N_\ell}^2 + 276S_{1,0}S_{1,4}{N_\ell}^2 \\
& + 1656S_{1,1}^2S_{2,0}{N_\ell}^2 + 6S_{0,4}S_{2,0}{N_\ell}^2 - 252S_{1,0}S_{1,2}S_{2,0}{N_\ell}^2 + 72S_{1,0}S_{1,1}S_{2,1}{N_\ell}^2 + 792S_{1,0}^2S_{2,2}{N_\ell}^2 \\
& - 306S_{2,0}S_{2,2}{N_\ell}^2 + 132S_{2,4}{N_\ell}^2 - 36S_{1,2}S_{3,0}{N_\ell}^2 - 288S_{1,1}S_{3,1}{N_\ell}^2 + 828S_{1,0}S_{3,2}{N_\ell}^2 \\
& - 144S_{0,4}S_{1,0}^2N_\ell - 1728S_{1,0}^2S_{1,1}^2N_\ell - 360S_{1,2}^2N_\ell - 540S_{2,1}^2N_\ell - 8S_{0,6}N_\ell - 360S_{1,0}^3S_{1,2}N_\ell \\
& + 672S_{1,1}S_{1,3}N_\ell - 264S_{1,0}S_{1,4}N_\ell - 2016S_{1,1}^2S_{2,0}N_\ell + 84S_{0,4}S_{2,0}N_\ell + 720S_{1,0}S_{1,2}S_{2,0}N_\ell \\
& + 2160S_{1,0}S_{1,1}S_{2,1}N_\ell - 936S_{1,0}^2S_{2,2}N_\ell + 756S_{2,0}S_{2,2}N_\ell - 48S_{2,4}N_\ell - 360S_{1,2}S_{3,0}N_\ell \\
& + 1008S_{1,1}S_{3,1}N_\ell - 792S_{1,0}S_{3,2}N_\ell + 9(N_\ell-2)(N_\ell-1)^2((N_\ell-1)N_\ell+4)S_{4,2}N_\ell \\
& + 3(N_\ell-2)(N_\ell-1)(3(N_\ell-5)N_\ell+20)S_{0,2}^3 \\
& - (N_\ell(N_\ell(N_\ell(N_\ell(N_\ell+4)-41)+40)+100)-80)S_{0,3}^2 + 48S_{0,4}S_{1,0}^2 - 144S_{1,0}^2S_{1,1}^2 \\
& + 288S_{1,2}^2 + 432S_{2,1}^2 + 144S_{1,0}^3S_{1,2} - 384S_{1,1}S_{1,3} + 96S_{1,0}S_{1,4} + 864S_{1,1}^2S_{2,0} - 48S_{0,4}S_{2,0} \\
& - 432S_{1,0}S_{1,2}S_{2,0} + 18(N_\ell-2)(N_\ell-1)S_{0,2}^2((N_\ell-4)(N_\ell-1)S_{2,0} - 2(N_\ell-3)S_{1,0}^2) \\
& - 864S_{1,0}S_{1,1}S_{2,1} + 6S_{0,3}(4(N_\ell(N_\ell+3)((N_\ell-6)N_\ell+10)-8)S_{1,0}S_{1,1} \\
& - (N_\ell((N_\ell-2)N_\ell({N_\ell}^2-17)+40)-32)S_{2,1}) + 432S_{1,0}^2S_{2,2} - 432S_{2,0}S_{2,2} + 288S_{1,2}S_{3,0} \\
& - 576S_{1,1}S_{3,1} + 288S_{1,0}S_{3,2} - 3(N_\ell-2)(N_\ell-1)S_{0,2}((N_\ell(N_\ell(2N_\ell-5)-5)+20)S_{0,4} \\
& + 3(-4S_{1,0}^4 + 4(N_\ell+1)S_{2,0}S_{1,0}^2 + 4((-{N_\ell}^2+N_\ell+4)S_{1,2} - 2(N_\ell-1)S_{3,0})S_{1,0} \\
& - ((N_\ell-5)N_\ell+12)S_{2,0}^2 + 2N_\ell((N_\ell-3)N_\ell-2)S_{2,2} + (N_\ell-1)N_\ell S_{4,0} + 4(-((N_\ell-5)N_\ell+8)S_{1,1}^2 \\
& + 4S_{2,2}+S_{4,0}))))\}
\end{aligned}
$$

For the sake of notation, we do not present the formula for the case $p = 4$ here, instead we refer the interested reader to our implementation available in the accompanying Python library `cmlmc-py`.

References

[1] Mark Ainsworth and J Tinsley Oden, *A posteriori error estimation in finite element analysis*, Vol. 37, John Wiley & Sons, 2011.

[2] C. Bierig and A. Chernov, *Convergence analysis of multilevel Monte Carlo variance estimators and application for random obstacle problems*, Numer. Math. **130** (2015), no. 4, 579–613.

[3] C. Bierig and A. Chernov, *Estimation of arbitrary order central statistical moments by the multilevel Monte Carlo method*, Stoch. Partial Differ. Equ. Anal. Comput. **4** (2016), no. 1, 3–40.

[4] Robert C Blattberg and Nicholas J Gonedes, *A comparison of the stable and student distributions as statistical models for stock prices*, The journal of business **47** (1974), no. 2, 244–280.

[5] Bo Young Chang, Peter Christoffersen, and Kris Jacobs, *Market skewness risk and the cross section of stock returns*, Journal of Financial Economics **107** (2013), no. 1, 46–68.

[6] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup, *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci. **14** (2011), no. 1, 3–15.

[7] Mark Drela, *A user's guide to MSES 3.05*, Massachusetts Institute of Technology (MIT), Cambridge (2007).

[8] P. S. Dwyer, *Moments of any rational integral isobaric sample moment function*, Ann. Math. Stat. **8** (1937), no. 1, 21–65.

[9] M. B. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res. **56** (2008), no. 3, 607–617.

[10] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numer. **24** (2015), 259–328.

[11] Werner Haase, Frans Brandsma, Eberhard Elsholz, Michael Leschziner, and Dieter Schwamborn, *EUROVAL - an european initiative on validation of CFD codes: Results of the EC/BRITE-EURAM project EUROVAL, 1990-1992*, Vol. 42, Springer-Verlag, 2013.

[12] P. R. Halmos, *The theory of unbiased estimation*, Ann. Math. Statist. **17** (1946), no. 1, 34–43.

[13] Alan Kraus and Robert H Litzenberger, *Skewness preference and the valuation of risk assets*, The Journal of Finance **31** (1976), no. 4, 1085–1100.

[14] Benoit Mandelbrot, *The variation of certain speculative prices*, The journal of business **36** (1963), no. 4, 394–419.

[15] Gamini Premaratne and Anil K Bera, *Modeling asymmetry and excess kurtosis in stock return data* (2000).

[16] C. Rose and M. D. Smith, *Mathematical statistics with Mathematica®*, Springer, 2002.

[17] J Clay Singleton and John Wingender, *Skewness persistence in common stock returns*, Journal of Financial and Quantitative Analysis **21** (1986), no. 3, 335–341.

[18] H. Sobieczky, *Parametric Airfoils and Wings*, Notes on Numerical Fluid Mechanics, edited by K. Fujii and G.S. Dulikravich **68** (1998), 71–88.

[19] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numer. Math. **125** (2013), no. 3, 569–600.

[20] V.A., *EXPERIMENTAL DATA BASE FOR COMPUTER PROGRAM ASSESSMENT - Report of the Fluid Dynamics Panel Working Group*, AGARD-AR-138 (1979).

[21] Rüdiger Verfürth, *A posteriori error estimates for nonlinear problems. finite element discretizations of elliptic equations*, Mathematics of Computation **62** (1994), no. 206, 445–475.

[22] D. L. Wallace, *Asymptotic approximations to distributions*, Ann. Math. Statist. **29** (1958), 635–654.

Recent publications:

INSTITUTE of MATHEMATICS
MATHICSE Group

Ecole Polytechnique Fédérale (EPFL)

CH-1015 Lausanne

2017

11.2017 ANDREAS FROMMER, CLAUDIA SCHIMMEL, MARCEL SCHWEITZER:
Bounds for the decay of the entries in inverses and Cauchy-Stieltjes functions of sparse, normal matrices

12.2017 SEBASTIAN KRUMSCHEID, FABIO NOBILE:
Multilevel Monte Carlo approximation of functions,

13.2017 R.N. SIMPSON, Z. LIU, R. VÁZQUEZ, J.A. EVANS:
An isogeometric boundary element method for electromagnetic scattering with compatible B-spline discretization

14.2017 NICCOLO DAL SANTO, SIMONE DEPARIS, ANDREA MANZONI:
A numerical investigation of multi space reduced basis preconditioners for parametrized elliptic advection-diffusion

15.2017 ASSYR ABDULLE, TIMOTHÉE POUCHON:
Effective models for long time wave propagation in locally periodic media

16.2017 ASSYR ABDULLE, MARCUS J. GROTE, ORANE JECKER:
FE-HMM for elastic waves in heterogeneous media

17.2017 JOHN A. EVANS, MICHAEL A. SCOTT, KENDRICK SHEPHERD, DEREK THOMAS, RAFAEL VÁZQUEZ:
Hierarchical B-spline complexes of discrete differential forms

18.2017 ELEONORA MUSHARBASH, FABIO NOBILE:
Symplectic dynamical low rank approximation of wave equations with random parameters

19.2017 ASSYR ABDULLE, IBRAHIM ALMUSLIMANI, GILLES VILMART:
Optimal explicit stabilized integrator of weak order one for stiff and ergodic stochastic differential equations

20.2017 ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, RAÚL TEMPONE, SÖREN WOLFERS:
Multilevel weighted least squares polynomial approximation

21.2017 NICCOLO DAL SANTO, SIMONE DEPARIS, ANDREA MANZONI, ALFIO QUARTERONI:
An algebraic least squares reduced basis method for the solution of parametrized Stokes equations

22.2017 ALEXEY CHERNOY, HAKON HOEL, KODY J. H. LAW, FABIO NOBILE, RAUL TEMPONE:
Multilevel ensemble Kalman filtering fors patio-temporal processes

23.2017 MICHELE PISARONI, SEBASTIAN KRUMSCHEID, FABIO NOBILE:
Quantifying uncertain system outputs via the multilevel Monte Carlo method – Part I: Central moment estimation

***