Journal of Computational Physics 341 (2017) 121–139

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

A low-rank control variate for multilevel Monte Carlo simulation of high-dimensional uncertain systems

Hillary R. Fairbanks [a], Alireza Doostan [b,*], Christian Ketelsen [c], Gianluca Iaccarino [d]

[a] Applied Mathematics and Statistics Department, University of Colorado, Boulder, CO 80309, USA
[b] Aerospace Engineering Sciences Department, University of Colorado, Boulder, CO 80309, USA
[c] Computer Science Department, University of Colorado, Boulder, CO 80309, USA
[d] Mechanical Engineering Department, Stanford University, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 31 October 2016
Received in revised form 27 March 2017
Accepted 28 March 2017
Available online 31 March 2017

Keywords:
Uncertainty quantification
Stochastic PDEs
Multilevel Monte Carlo
Control variate
Low-rank approximation
Multifidelity
Interpolative decomposition

ABSTRACT

Multilevel Monte Carlo (MLMC) is a recently proposed variation of Monte Carlo (MC) simulation that achieves variance reduction by simulating the governing equations on a series of spatial (or temporal) grids with increasing resolution. Instead of directly employing the fine grid solutions, MLMC estimates the expectation of the quantity of interest from the coarsest grid solutions as well as differences between each two consecutive grid solutions. When the differences corresponding to finer grids become smaller, hence less variable, fewer MC realizations of finer grid solutions are needed to compute the difference expectations, thus leading to a reduction in the overall work. This paper presents an extension of MLMC, referred to as *multilevel control variates (MLCV)*, where a low-rank approximation to the solution on each grid, obtained primarily based on coarser grid solutions, is used as a control variate for estimating the expectations involved in MLMC. Cost estimates as well as numerical examples are presented to demonstrate the advantage of this new MLCV approach over the standard MLMC when the solution of interest admits a low-rank approximation and the cost of simulating finer grids grows fast.

1. Introduction

The use of uncertainty quantification as a tool to assess the prediction accuracy of simulation models of physical systems has been increasing at a rapid rate over the last decade. By accounting for the uncertainties of input data in models, such as initial conditions, boundary conditions, or other physical parameters, the objective is to establish the predictive capabilities of simulations by quantifying the uncertainty in the quantities of interest (QoI's). To this end and within the probabilistic framework, several methods, e.g., polynomial chaos expansions [1–3] and stochastic collocation [4,5], have been developed and proven successful in various applications. However, it is known that the computational cost of these methods grows rapidly as a function of the number of random variables describing model uncertainties, a phenomenon referred to as *curse of dimensionality*.

An alternative class of techniques rely on the Monte Carlo (MC) simulation or its variants, where the statistics of the QoI are estimated using an ensemble of (random) realizations of the QoI. The cost of such estimations, while may be

* Corresponding author.
E-mail address: alireza.doostan@colorado.edu (A. Doostan).

http://dx.doi.org/10.1016/j.jcp.2017.03.060
0021-9991/

prohibitive, is *formally* independent of the number of input variables. In details, let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d)$ denote the d-vector of random variables, with joint probability density function $\rho_{\boldsymbol{\xi}}(\boldsymbol{\xi})$, representing the uncertainty in the inputs. Let $Q = Q(\boldsymbol{\xi})$ denote a scalar-valued QoI depending on $\boldsymbol{\xi}$ and Q_M its approximation obtained via simulation. The subscript M denotes the number of deterministic degrees of freedom, e.g., number of grid points in a finite element model, controlling the accuracy of Q_M relative to Q. The goal is to approximate the statistics of Q, e.g., the mean of Q, $\mathbb{E}[Q]$, using the realizations of Q_M. Given a set of N samples of inputs, each denoted by $\boldsymbol{\xi}^{(i)}$ and drawn according to $\rho_{\boldsymbol{\xi}}(\boldsymbol{\xi})$, and the corresponding realizations of Q_M, given by $Q_M^{(i)} = Q_M(\boldsymbol{\xi}^{(i)})$, the MC approximation of $\mathbb{E}[Q]$ is

$$\mathbb{E}[Q] \approx \mathbb{E}[Q_M] \approx \hat{Q}_{M,N}^{MC} = \frac{1}{N}\sum_{i=1}^{N} Q_M^{(i)}. \tag{1}$$

Following the notation in [6], $\hat{Q}_{M,N}^{MC}$ in (1) is the MC estimator of $\mathbb{E}[Q_M]$ using N samples of Q_M with the Mean Square Error (MSE)

$$\mathrm{MSE}(\hat{Q}_{M,N}^{MC}, \mathbb{E}[Q]) = \frac{1}{N}\mathbb{V}[Q_M] + (\mathbb{E}[Q_M - Q])^2, \tag{2}$$

where $\mathbb{V}$ is the variance operator and $\mathrm{MSE}(\hat{Q}_{M,N}^{MC}, \mathbb{E}[Q])$ denotes the MSE of $\hat{Q}_{M,N}^{MC}$ with respect to $\mathbb{E}[Q]$. We note that, in this paper, the hat operator indicates the MC estimator of the corresponding expectation. In (2), the MSE is decomposed into the *sampling error* $\frac{1}{N}\mathbb{V}[Q_M]$, controlled by the variance of Q_M and the number of samples, and the *discretization error* $(\mathbb{E}[Q_M - Q])^2$, which measures how closely the model simulates the true solution. As can be seen from (2), the sampling error decays slowly as a function of N, but with a rate that is independent of the dimension d, implying that the standard MC simulation does not formally suffer from the curse of dimensionality.

Aside from the necessary refinement in the model to reduce the discretization error, there are only two options to improve the MSE of an MC estimate: increasing the sample size N or using a variance reduction technique. Due to the cost incurred by the first option, it is more practical to consider the use of a variance reduction technique, such as importance sampling or control variates (CV) [7]. In particular, the CV approach considers a second quantity Z, such that it is correlated with Q_M, is cheaper than Q_M to evaluate, and whose expectation is either known or can be approximated with relatively small cost. Then a new variable,

$$W = Q_M - \theta(Z - \mathbb{E}[Z]),$$

is constructed that has the same mean as Q_M, i.e., $\mathbb{E}[W] = \mathbb{E}[Q_M]$, thus suggesting the use of MC estimate of $\mathbb{E}[W]$ as a proxy for $\mathbb{E}[Q_M]$. In doing so, the gain is that, depending on the choice of θ, the MC estimator of $\mathbb{E}[W]$ features a reduced MSE (or variance). Stated differently, a smaller number of W realizations, hence Q_M realizations, are needed for a comparable MSE when CV is applied. For scenarios when Z is poorly correlated with Q_M, a notable MSE reduction is not observed. If, in addition, the cost of estimating $\mathbb{E}[Z]$ is large, it is likely that this CV will not result in a cost improvement over standard MC.

Multilevel Monte Carlo (MLMC), proposed in [8,9], is a generalization of CV, which constructs a sequence of control variates Z based on approximations of Q on a set of models that are cheaper to simulate than the one for Q_M, hence the term *multilevel*. A common example of a cheaper model is to approximate Q on coarser grids with number of degrees of freedom smaller than M. While the notion of levels can go beyond a grid-based construction, we limit the scope of this study to such an approach. For the interest of a simpler introduction, we delay the full presentation of MLMC to Section 2, and instead focus on the two-level formulation next.

Taking $\theta = 1$ and $Z = Q_m$, with $m < M$, to be the QoI approximated from a coarser grid than that of Q_M, the two-level MLMC variable is given by $W = Q_M - (Q_m - \mathbb{E}[Q_m]) = \mathbb{E}[Q_m] + (Q_M - Q_m)$, with expected value

$$\begin{aligned}\mathbb{E}[W] &= \mathbb{E}[Q_m] + \mathbb{E}[Q_M - Q_m] \\ &= \mathbb{E}[Q_M].\end{aligned} \tag{3}$$

To approximate $\mathbb{E}[W]$, or equivalently $\mathbb{E}[Q_M]$, MC is applied independently to the two expectations in the right-hand-side of the first equation in (3),

$$\hat{W} = \frac{1}{N_m}\sum_{i=1}^{N_m} Q_m^{(i)} + \frac{1}{N_M}\sum_{i=1}^{N_M}(Q_M^{(i)} - Q_m^{(i)}). \tag{4}$$

As compared to the standard MC estimator of $\mathbb{E}[Q_M]$ given in (1), the estimation of $\mathbb{E}[Q_m]$ in (4) also involves drawing samples of Q_m, which are less expensive. More importantly, when Q_m is close to Q_M, estimating $\mathbb{E}[Q_M - Q_m]$ requires fewer samples of Q_M, as $(Q_M - Q_m)$ features a smaller variance. In practice, depending on the cost of simulating the two models as well as the variances of Q_m and $(Q_M - Q_m)$, the numbers of samples of Q_m, N_m, and Q_M, N_M, are selected such that the overall estimation cost, for a given accuracy, is minimal. MLMC expands upon this concept by including multiple levels, as delineated in Section 2.